

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
22 February 2001 (22.02.2001)

PCT

(10) International Publication Number
WO 01/12662 A2

(51) International Patent Classification⁷: C07K 14/00

(21) International Application Number: PCT/US00/22315

(22) International Filing Date: 14 August 2000 (14.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/149,641 17 August 1999 (17.08.1999) US
60/164,203 9 November 1999 (09.11.1999) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:
US 60/149,641 (CIP)
Filed on 17 August 1999 (17.08.1999)
US 60/164,203 (CIP)
Filed on 9 November 1999 (09.11.1999)

(71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). BURFORD, Neil [GB/US]; 1308 4th Avenue, San Francisco, CA 94122 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/12662 A2

(54) Title: MEMBRANE ASSOCIATED PROTEINS

(57) Abstract: The invention provides human membrane associated proteins (MEMAP) and polynucleotides which identify and encode MEMAP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of MEMAP.

MEMBRANE ASSOCIATED PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of membrane associated proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, neurological and gastrointestinal disorders.

BACKGROUND OF THE INVENTION

Eukaryotic cells are surrounded by plasma membranes which enclose the cell and maintain an environment inside the cell that is distinct from its surroundings. In addition, eukaryotic organisms are distinct from prokaryotes in possessing many intracellular organelle and vesicle structures. Many of the metabolic reactions which distinguish eukaryotic biochemistry from prokaryotic biochemistry take place within these structures. The plasma membrane and the membranes surrounding organelles and vesicles are composed of phosphoglycerides, fatty acids, cholesterol, phospholipids, glycolipids, proteoglycans, and proteins. These components confer identity and functionality to the membranes with which they associate.

Integral Membrane Proteins

The majority of known integral membrane proteins are transmembrane proteins (TM) which are characterized by an extracellular, a transmembrane, and an intracellular domain. TM domains are typically comprised of 15 to 25 hydrophobic amino acids which are predicted to adopt an α -helical conformation. TM proteins are classified as bitopic (Types I and II) and polytopic (Types III and IV) (Singer, S.J. (1990) *Ann. Rev. Cell Biol.* 6:247-96). Bitopic proteins span the membrane once while polytopic proteins contain multiple membrane-spanning segments. TM proteins that act as cell-surface receptor proteins involved in signal transduction include growth and differentiation factor receptors, and receptor-interacting proteins such as *Drosophila* pecanex and frizzled proteins, LIV-1 protein, NF2 protein, and GNS1/SUR4 eukaryotic integral membrane proteins. TM proteins also act as transporters of ions or metabolites, such as gap junction channels (connexins) and ion channels, and as cell anchoring proteins, such as lectins, integrins, and fibronectins. TM proteins act as vesicle organelle-forming molecules, such as calveolins, or as cell recognition molecules, such as cluster of differentiation (CD) antigens, glycoproteins, and mucins.

Many membrane proteins (MPs) contain amino acid sequence motifs that target these proteins to specific subcellular sites. Examples of these motifs include PDZ domains, KDEL, RGD, NGR, and GSL sequence motifs, von Willebrand factor A (vWFA) domains, and EGF-like domains. RGD, NGR, and GSL motif-containing peptides have been used as drug delivery agents in cancer

treatments which target tumor vasculature (Arap, W. et al. (1998) *Science*, 279:377-380). Furthermore, MPs may also contain amino acid sequence motifs, such as the carbohydrate recognition domain (CRD), also known as the C-type lectin domain, that mediate interactions with extracellular or intracellular molecules.

5 Membrane proteins may also interact with and regulate the properties of the membrane lipids. Phospholipid scramblase, a type II plasma membrane protein, mediates calcium dependent movement of phospholipids (PL) between membrane leaflets. Calcium induced remodeling of plasma membrane PL plays a key role in expression of platelet anticoagulant activity and in clearance of injured or apoptotic cells (Zhou Q. et al. (1997) *J. Biol. Chem.* 272:18240-18244). Scott syndrome, a 10 bleeding disorder, is caused by an inherited deficiency in plasma membrane PL scramblase function (Online Mendelian Inheritance in Man (OMIM) *262890 Platelet Receptor for Factor X, Deficiency of).

15 Chemical modification of amino acid residue side chains alters the manner in which MPs interact with other molecules, such as phospholipid membranes. Examples of such chemical modifications to amino acid residue side chains are covalent bond formation with glycosaminoglycans, oligosaccharides, phospholipids, acetyl and palmitoyl moieties, ADP-ribose, phosphate, and sulphate groups.

20 One function of TM proteins is to facilitate cell-cell communication. The slit proteins are extracellular matrix proteins expressed by cells at the ventral midline of the nervous system. Slit proteins are ligands for the repulsive guidance receptor Robo and thus play a role in repulsive axon guidance (Brose, K. et al. (1999) *Cell* 96:795-806).

25 In some cases TM proteins serve as transporters or channels in the cell membrane. For example, the mouse transporter protein (MTP) has four transmembrane domains and resides in an intracellular membrane compartment. MTP can mediate transport of nucleosides in vitro. The role of MTP in the cell may therefore be to transfer nucleosides between the cytosol and the lumen of intracellular organelles (Hogue, D. L. (1996) *J. Biol. Chem.* 271:9801-9808). The human stomatin-like protein (hSLP-1), expressed primarily in the brain, contains an N-terminal domain similar to the erythrocyte internal membrane protein stomatin, as well as a non-specific lipid transfer protein domain at the C-terminus. hSLP-1 is the human homologue of the *C. elegans* behavioral gene unc-30, which is believed to be involved in lipid transfer between closely apposed membranes (Seidel, G. and Prohaska, R (1998) *Gene* 225:23-29).

30 The transmembrane 4 superfamily (TM4SF) or tetraspan family is a multigene family encoding type III integral membrane proteins (Wright, M.D. and Tomlinson, M.G. (1994) *Immunol. Today* 15:588-594). TM4SF is comprised of membrane proteins which traverse the cell membrane 35 four times. Members of the TM4SF include platelet and endothelial cell membrane proteins,

melanoma-associated antigens, leukocyte surface glycoproteins, colonial carcinoma antigens, tumor-associated antigens, and surface proteins of the schistosome parasites (Jankowski, S.A. (1994) *Oncogene* 9:1205-1211). Members of the TM4SF share about 25-30% amino acid sequence identity with one another.

5 A number of TM4SF members have been implicated in signal transduction, control of cell adhesion, regulation of cell growth and proliferation, including development and oncogenesis, and cell motility, including tumor cell metastasis. Expression of TM4SF proteins is associated with a variety of tumors and the level of expression may be altered when cells are growing or activated.

Tumor antigens are cell surface molecules that are differentially expressed in tumor cells
10 relative to normal cells. Tumor antigens distinguish tumor cells immunologically from normal cells and provide diagnostic and therapeutic targets for human cancers (Takagi, S. et al. (1995) *Int. J. Cancer* 61: 706-715; Liu, E. et al. (1992) *Oncogene* 7: 1027-1032). For example, the biliary glycoprotein-encoding gene is a member of the human carcinoembryonic antigen family, which are important tumor markers for colorectal carcinomas (Hammarstrom, S. (1999) *Semin. Cancer Bio.* 15: 9:67-81). Another example is the neuron and testis specific protein Ma1, a marker for paraneoplastic neuronal disorders (Dalmau, J. et al. (1999) *Brain* 122:27-39).

Other types of cell surface antigens include those identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs
20 directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and
25 verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI).
(Reviewed in Barclay, A. N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

30 The TM cell surface glycoprotein CD69 is an early activation antigen of T lymphocytes. CD69 is homologous to members of a supergene family of type II integral membrane proteins having C-type lectin domains. Although the precise functions of the CD-69 antigen is not known, evidence suggests that these proteins transmit mitogenic signals across the plasma membrane and are up-regulated in response to lymphocyte activation (Hamann, J. et. al. (1993) *J. Immunol.* 150:4920-35 4927).

Macrophages are involved in functions including clearance of senescent or apoptotic cells, cytokine production, hemopoiesis, bone resorption, antigen transport, and neuroendocrine regulation. These diverse roles are influenced by specialized macrophage plasma membrane proteins. The murine macrophage restricted C-type lectin is a type II integral membrane protein expressed 5 exclusively in macrophages. The strong expression of this protein in bone marrow suggests a hemopoietic function, while the lectin domain suggests it may be involved in cell-cell recognition (Balch, S. G. et al. (1998) *J. Biol. Chem.* 273:18656-18664).

The surface of red blood cells is populated with characteristic glycoproteins, such as the major sialoglycoproteins glycophorin A and B. Red blood cells lacking either glycophorin A or B are 10 resistant to infection with the malaria parasite *Plasmodium falciparum* (OMIM Entry 111300 Blood Group-MN Locus). White blood cells also possess characteristic surface glycoproteins, such as the plasma cell glycoprotein-1 (PC-1). PC-1 is expressed on the surface of plasma cells, which are terminally differentiated, antibody-secreting B-lymphocytes. The extracellular domain of PC-1 has 15 nucleotide phosphodiesterase (pyrophosphatase) activity (Funakoshi, I. et al. (1992) *Arch. Biochem. Biophys.* 295:180-187). Phosphodiesterase activity is associated with the hydrolytic removal of nucleotide subunits from oligonucleotides. Although the precise physiological role of PC-1 is not clear, increased PC-1 phosphodiesterase activity has been correlated with insulin resistance in patients with noninsulin-dependent diabetes mellitus, with abnormalities of bone mineralization and calcification, and with defects in renal tubule function. In addition, it appears that hPC-1 and mPC-1 20 are members of a multigene family of transmembrane phosphodiesterases with extracellular active sites. These enzymes may play a role in regulating the concentration of pharmacologically active extracellular compounds such as adenosine or other nucleotide derivatives in a variety of tissues and cell types. (Reviewed in Goding, J. W. et al. (1998) *Immunol. Rev.* 161:11-26.)

25 Peripheral and Anchored Membrane Proteins

Some membrane proteins are not membrane-spanning but are attached to the plasma membrane via membrane anchors or interactions with integral membrane proteins. Membrane anchors are covalently joined to a protein post-translationally and include such moieties as prenyl, myristyl, and glycosylphosphatidyl inositol (GPI) groups. Membrane localization of peripheral and 30 anchored proteins is important for their function in processes such as receptor-mediated signal transduction. For example, prenylation of Ras is required for its localization to the plasma membrane and for its normal and oncogenic functions in signal transduction.

The pancortins are a group of four glycoproteins which are predominantly expressed in the cerebral cortex of adult rodents. Immunological localization indicates that the pancortins are 35 endoplasmic reticulum anchored proteins. The pancortins share a common sequence in the middle of

their structure, but have alternative sequences at both ends due to differential promoter usage and alternative splicing. Each pancortin appears to be differentially expressed and may perform different functions in the brain (Nagano, T. et al. (1998) Mol. Brain Res. 53:13-23).

The discovery of new membrane associated proteins and the polynucleotides encoding them
5 satisfies a need in the art by providing new compositions which are useful in the diagnosis,
prevention, and treatment of cell proliferative, autoimmune/inflammatory, neurological and
gastrointestinal disorders.

SUMMARY OF THE INVENTION

10 The invention features purified polypeptides, membrane associated proteins, referred to
collectively as "MEMAP" and individually as "MEMAP-1," "MEMAP-2," "MEMAP-3," "MEMAP-
4," "MEMAP-5," "MEMAP-6," "MEMAP-7," "MEMAP-8," "MEMAP-9," "MEMAP-10,"
"MEMAP-11," "MEMAP-12," "MEMAP-13," "MEMAP-14," "MEMAP-15," "MEMAP-16,"
"MEMAP-17," "MEMAP-18," "MEMAP-19," "MEMAP-20," "MEMAP-21," "MEMAP-22,"
15 "MEMAP-23," "MEMAP-24," "MEMAP-25," "MEMAP-26," "MEMAP-27," "MEMAP-28,"
"MEMAP-29," "MEMAP-30," "MEMAP-31," "MEMAP-32," "MEMAP-33," "MEMAP-34,"
"MEMAP-35," "MEMAP-36," and "MEMAP-37." In one aspect, the invention provides an isolated
polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino
acid sequence selected from the group consisting of SEQ ID NO:1-37, b) a naturally occurring amino
20 acid sequence having at least 90% sequence identity to an amino acid sequence selected from the
group consisting of SEQ ID NO:1-37, c) a biologically active fragment of an amino acid sequence
selected from the group consisting of SEQ ID NO:1-37, and d) an immunogenic fragment of an amino
acid sequence selected from the group consisting of SEQ ID NO:1-37. In one alternative, the
invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-37.

25 The invention further provides an isolated polynucleotide encoding a polypeptide comprising
an amino acid sequence selected from the group consisting of a) an amino acid sequence selected
from the group consisting of SEQ ID NO:1-37, b) a naturally occurring amino acid sequence having
at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ
ID NO:1-37, c) a biologically active fragment of an amino acid sequence selected from the group
30 consisting of SEQ ID NO:1-37, and d) an immunogenic fragment of an amino acid sequence selected
from the group consisting of SEQ ID NO:1-37. In one alternative, the polynucleotide encodes a
polypeptide selected from the group consisting of SEQ ID NO:1-37. In another alternative, the
polynucleotide is selected from the group consisting of SEQ ID NO:38-74.

35 Additionally, the invention provides a recombinant polynucleotide comprising a promoter
sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid

sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:38-74, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:38-74, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group

consisting of SEQ ID NO:38-74, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:38-74, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the

5 sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the

10 probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:38-74, b) a naturally occurring polynucleotide sequence having at least 15 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:38-74, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, 20 optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the 25 group consisting of SEQ ID NO:1-37, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-37. The invention additionally provides a 30 method of treating a disease or condition associated with decreased expression of functional MEMAP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, b) a naturally 35 occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence

selected from the group consisting of SEQ ID NO:1-37, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting 5 agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional MEMAP, comprising administering to a patient in need of such treatment the composition.

10 Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, and d) an immunogenic 15 fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with 20 overexpression of functional MEMAP, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37. The method comprises a) 25 combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a) 30 an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, b) a naturally

occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-37. The 5 method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a 10 compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:38-74, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered 15 expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide comprising a polynucleotide sequence selected from the 20 group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:38-74, ii) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:38-74, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific 25 hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:38-74, ii) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:38-74, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence 30 complementary to ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of the above polynucleotide sequence; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is 35 indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding MEMAP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of MEMAP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding MEMAP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the

invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

“MEMAP” refers to the amino acid sequences of substantially purified MEMAP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, 5 and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term “agonist” refers to a molecule which intensifies or mimics the biological activity of MEMAP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of MEMAP either by directly interacting with MEMAP or by acting on components of the biological pathway in which MEMAP participates.

10 An “allelic variant” is an alternative form of the gene encoding MEMAP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give 15 rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

“Altered” nucleic acid sequences encoding MEMAP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as MEMAP or a polypeptide with at least one functional characteristic of MEMAP. Included within this definition 20 are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding MEMAP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding MEMAP. The encoded protein may also be “altered,” and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a 25 functionally equivalent MEMAP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of MEMAP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains 30 having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms “amino acid” and “amino acid sequence” refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic 35 molecules. Where “amino acid sequence” is recited to refer to a sequence of a naturally occurring

protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of MEMAP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of MEMAP either by directly interacting with MEMAP or by acting on components of the biological pathway in which

MEMAP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind MEMAP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the

designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic MEMAP, or of any oligopeptide

5 thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

10 A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding MEMAP or fragments of MEMAP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be
15 associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (PE Biosystems, 20 Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

25 "Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

30

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
35	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His

	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
5	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
10	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

20 The term "derivative" refers to a chemically modified polynucleotide or polypeptide.

Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule.

25 A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

30 A "fragment" is a unique portion of MEMAP or the polynucleotide encoding MEMAP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the

present embodiments.

1 A fragment of SEQ ID NO:38-74 comprises a region of unique polynucleotide sequence that
specifically identifies SEQ ID NO:38-74, for example, as distinct from any other sequence in the
genome from which the fragment was obtained. A fragment of SEQ ID NO:38-74 is useful, for
5 example, in hybridization and amplification technologies and in analogous methods that distinguish
SEQ ID NO:38-74 from related polynucleotide sequences. The precise length of a fragment of SEQ
ID NO:38-74 and the region of SEQ ID NO:38-74 to which the fragment corresponds are routinely
determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

10 A fragment of SEQ ID NO:1-37 is encoded by a fragment of SEQ ID NO:38-74. A fragment
of SEQ ID NO:1-37 comprises a region of unique amino acid sequence that specifically identifies
SEQ ID NO:1-37. For example, a fragment of SEQ ID NO:1-37 is useful as an immunogenic peptide
for the development of antibodies that specifically recognize SEQ ID NO:1-37. The precise length of
a fragment of SEQ ID NO:1-37 and the region of SEQ ID NO:1-37 to which the fragment
corresponds are routinely determinable by one of ordinary skill in the art based on the intended
15 purpose for the fragment.

A “full-length” polynucleotide sequence is one containing at least a translation initiation
codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A
“full-length” polynucleotide sequence encodes a “full-length” polypeptide sequence.

20 “Homology” refers to sequence similarity or, interchangeably, sequence identity, between
two or more polynucleotide sequences or two or more polypeptide sequences.

25 The terms “percent identity” and “% identity,” as applied to polynucleotide sequences, refer
to the percentage of residue matches between at least two polynucleotide sequences aligned using a
standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps
in the sequences being compared in order to optimize alignment between two sequences, and
therefore achieve a more meaningful comparison of the two sequences.

30 Percent identity between polynucleotide sequences may be determined using the default
parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e
sequence alignment program. This program is part of the LASERGENE software package, a suite of
molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in
Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS
8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as
follows: Ktuple=2, gap penalty=5, window=4, and “diagonals saved”=4. The “weighted” residue
weight table is selected as the default. Percent identity is reported by CLUSTAL V as the “percent
similarity” between aligned polynucleotide sequences.

35 Alternatively, a suite of commonly used and freely available sequence comparison algorithms

is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) *J. Mol. Biol.* 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

15 *Penalty for mismatch: -2*

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

20 *Filter: on*

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous 25 nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

30 Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

35 The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative

substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

15 *Open Gap: 11 and Extension Gap: 1 penalties*

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

20 Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for chromosome replication, segregation and maintenance.

30 The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity.

Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive 5 conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

10 Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and 15 conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present 20 invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents 25 include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid 30 sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

35 The words "insertion" and "addition" refer to changes in an amino acid or nucleotide

sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

“Immune response” can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect 5 cellular and systemic defense systems.

An “immunogenic fragment” is a polypeptide or oligopeptide fragment of MEMAP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term “immunogenic fragment” also includes any polypeptide or oligopeptide fragment of MEMAP which is useful in any of the antibody production methods disclosed herein or known in 10 the art.

The term “microarray” refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms “element” and “array element” refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

15 The term “modulate” refers to a change in the activity of MEMAP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of MEMAP.

20 The phrases “nucleic acid” and “nucleic acid sequence” refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the 25 antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

25 “Operably linked” refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

30 “Peptide nucleic acid” (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNA 35 preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

35 “Post-translational modification” of an MEMAP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of MEMAP.

"Probe" refers to nucleic acid sequences encoding MEMAP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

5 "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous 10 nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

15 Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs 20 can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 25 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from 30 megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection 35 programs may also be obtained from their respective sources and modified to meet the user's specific

needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both 5 unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

10 A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have 15 been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

20 Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

25 A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

30 "Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

35 The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding MEMAP, or fragments thereof, or MEMAP itself, may comprise a bodily fluid; an

extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or 5 synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

10 The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

15 A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

16 "Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

20 A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

25 "Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well 30 as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

35 A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with

a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants, and animals. The isolated DNA of the present invention can be 5 introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook, J. et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having 10 at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, 15 an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to 20 another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a 25 propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at 30 least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human membrane associated proteins (MEMAP), the polynucleotides encoding MEMAP, and the use of these compositions for the 35 diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, neurological and

gastrointestinal disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding MEMAP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each MEMAP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. In some cases, GenBank sequence identifiers are also shown in column 5. The Incyte clones and GenBank cDNA sequences, where indicated, in column 5 were used to assemble the consensus nucleotide sequence of each MEMAP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding MEMAP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:38-74 and to distinguish between SEQ ID NO:38-74 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express MEMAP as a fraction of total tissues expressing MEMAP. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing MEMAP as a fraction of total tissues expressing MEMAP. Column 5 lists the vectors used to subclone each cDNA library. Of particular note is the expression of SEQ ID NO:41, SEQ ID NO:48, and SEQ ID NO:56 in nervous tissues, of SEQ ID NO:52, SEQ ID NO:65, and SEQ ID NO:74 in gastrointestinal tissues, and of SEQ ID NO:55 in hematopoietic/immune tissues.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding MEMAP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:38 maps to chromosome 4 within the interval from 77.9 to 86.0 centiMorgans, to chromosome 6 within the interval from 132.7 to 144.4 centiMorgans, and to chromosome 14 within the interval from 89.4 to 103.7 centiMorgans. The interval on chromosome 4 from 77.9 to 86.0 centiMorgans also contains a gene associated with deoxycytidine kinase deficiency. The interval on 5 chromosome 6 from 132.7 to 144.4 centiMorgans also contains genes associated with peroxisomal disorders and leukemia. The interval on chromosome 14 from 89.4 to 103.7 centiMorgans also contains genes associated with spinocerebellar ataxia and protease inhibitor deficiencies. SEQ ID NO:39 maps to chromosome 2 within the interval from 236.2 to 269.5 centiMorgans, and to the X chromosome within the interval from 94.4 to 97.4 centiMorgans. The interval on chromosome 2 from 236.2 to 269.5 centiMorgans also contains genes associated with Crigler-Najjar syndrome, 10 Oguchi disease, and oxaolis I. The interval on the X chromosome from 94.4 to 97.4 centiMorgans also contains genes associated with Charcot-Marie tooth disease, X-linked severe combined immunodeficiency, alpha thalassemia/mental retardation syndrome, Menkes' syndrome, and choroideremia. SEQ ID NO:42 maps to chromosome 1 within the interval from 218.2 to 232.0 centiMorgans. This interval also contains genes associated with familial hypertrophic 15 cardiomyopathy, malignant hyperthermia, and hypokalemic periodic paralysis. SEQ ID NO:44 maps to chromosome 7 within the interval from 136.4 to 145.8 centiMorgans, to chromosome 14 within the interval from 28.0 to 32.9 centiMorgans, and to chromosome 14 within the interval from 71.5 to 73.7 centiMorgans. The interval on chromosome 7 from 136.4 to 145.8 centiMorgans also contains genes 20 associated with diphosphoglycerate mutase deficiency. SEQ ID NO:60 maps to chromosome 7 within the interval from 167.6 to 184.0 centiMorgans, and to chromosome 14 within the interval from 50.0 to 59.0 centiMorgans. SEQ ID NO:63 maps to chromosome 8 within the interval from 101.0 to 125.8 centiMorgans, and to chromosome 8 within the interval from 132.4 to 135.1 centiMorgans. SEQ ID NO:67 maps to chromosome 4 within the interval from 145.3 to 146.4 centiMorgans.

25 The invention also encompasses MEMAP variants. A preferred MEMAP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the MEMAP amino acid sequence, and which contains at least one functional or structural characteristic of MEMAP.

30 The invention also encompasses polynucleotides which encode MEMAP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:38-74, which encodes MEMAP. The polynucleotide sequences of SEQ ID NO:38-74, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

35 The invention also encompasses a variant of a polynucleotide sequence encoding MEMAP.

In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding MEMAP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:38-74 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:38-74. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of MEMAP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding MEMAP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring MEMAP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode MEMAP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring MEMAP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding MEMAP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding MEMAP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode MEMAP and MEMAP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding MEMAP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:38-74 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and 35 S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.*

152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (PE Biosystems, Foster City CA), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (PE Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (PE Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding MEMAP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National

Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been

5 size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze

10 the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, PE Biosystems), and the entire

15 process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode MEMAP may be cloned in recombinant DNA molecules that direct expression of

20 MEMAP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express MEMAP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter MEMAP-encoding sequences for a variety of purposes including,

25 but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

30 The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of MEMAP, such as its biological or enzymatic activity or its

35 ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of

gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" 5 breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable 10 manner.

In another embodiment, sequences encoding MEMAP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, MEMAP itself or a fragment thereof may be synthesized using chemical methods. For 15 example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (PE Biosystems). Additionally, the amino acid sequence of MEMAP, or any part thereof, may be altered during direct synthesis and/or 20 combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by 25 sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active MEMAP, the nucleotide sequences encoding MEMAP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, 30 constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding MEMAP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding MEMAP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding MEMAP and its initiation 35 codon and upstream regulatory sequences are inserted into the appropriate expression vector, no

additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic.

5 The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding MEMAP and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. (See, e.g., Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

15 A variety of expression vector/host systems may be utilized to contain and express sequences encoding MEMAP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, *supra*; Ausubel, *supra*; Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184; Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-25 1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses, 30 adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Buller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. (1994) *Mol. Immunol.* 31(3):219-226; and Verma, I.M. and N. Somia (1997) *Nature* 389:239-242.)

35 The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding MEMAP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding MEMAP can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSORT1 plasmid (Life Technologies). Ligation of sequences encoding MEMAP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of MEMAP are needed, e.g. for the production of antibodies, vectors which direct high level expression of MEMAP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of MEMAP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, *supra*; and Scorer, *supra*.)

Plant systems may also be used for expression of MEMAP. Transcription of sequences encoding MEMAP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, *supra*; Broglie, *supra*; and Winter, *supra*.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding MEMAP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses MEMAP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of

DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)

5 For long term production of recombinant proteins in mammalian systems, stable expression of MEMAP in cell lines is preferred. For example, sequences encoding MEMAP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in
10 enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These
15 include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) *Cell* 11:223-232; Lowy, I. et al. (1980) *Cell* 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat*
20 confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci. USA* 77:3567-3570; Colbere-Garapin, F. et al. (1981) *J. Mol. Biol.* 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins
25 (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system.
(See, e.g., Rhodes, C.A. (1995) *Methods Mol. Biol.* 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is
30 also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding MEMAP is inserted within a marker gene sequence, transformed cells containing sequences encoding MEMAP can be identified by the absence of marker gene function.

Alternatively, a marker gene can be placed in tandem with a sequence encoding MEMAP under the control of a single promoter. Expression of the marker gene in response to induction or selection
35 usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding MEMAP and that express MEMAP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

5 Immunological methods for detecting and measuring the expression of MEMAP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing 10 monoclonal antibodies reactive to two non-interfering epitopes on MEMAP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana 15 Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding MEMAP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. 20 Alternatively, the sequences encoding MEMAP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia 25 Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding MEMAP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein 30 produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode MEMAP may be designed to contain signal sequences which direct secretion of MEMAP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the 35 inserted sequences or to process the expressed protein in the desired fashion. Such modifications of

the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for 5 post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding MEMAP may be ligated to a heterologous sequence resulting in translation of a 10 fusion protein in any of the aforementioned host systems. For example, a chimeric MEMAP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of MEMAP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), 15 maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity 20 purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the MEMAP encoding sequence and the heterologous protein sequence, so that MEMAP may be cleaved away from the heterologous moiety following 25 purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled MEMAP may be achieved 30 in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S -methionine.

MEMAP of the present invention or fragments thereof may be used to screen for compounds 35 that specifically bind to MEMAP. At least one and up to a plurality of test compounds may be screened for specific binding to MEMAP. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of

MEMAP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which MEMAP binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the 5 compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express MEMAP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing MEMAP or cell membrane fractions which contain MEMAP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either MEMAP 10 or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with MEMAP, either in solution or affixed to a solid support, and detecting the binding of MEMAP to the compound. 15 Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

MEMAP of the present invention or fragments thereof may be used to screen for compounds 20 that modulate the activity of MEMAP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for MEMAP activity, wherein MEMAP is combined with at least one test compound, and the activity of MEMAP in the presence of a test compound is compared with the activity of MEMAP in the absence of the test compound. A change in the activity of MEMAP in the presence of the test compound is 25 indicative of a compound that modulates the activity of MEMAP. Alternatively, a test compound is combined with an in vitro or cell-free system comprising MEMAP under conditions suitable for MEMAP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of MEMAP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

30 In another embodiment, polynucleotides encoding MEMAP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse 35 embryo and grown in culture. The ES cells are transformed with a vector containing the gene of

interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic Acids Res.* 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

10 Polynucleotides encoding MEMAP may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. 15 (1998) *Science* 282:1145-1147).

10 Polynucleotides encoding MEMAP can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding MEMAP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and 20 the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress MEMAP, e.g., by secreting MEMAP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) *Biotechnol. Annu. Rev.* 4:55-74).

25 THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists 30 between regions of MEMAP and membrane associated proteins. In addition, the expression of MEMAP is closely associated with neurological and gastrointestinal tissues, cancer, cell proliferation, and inflammation/trauma. Therefore, MEMAP appears to play a role in cell proliferative, autoimmune/inflammatory, neurological and gastrointestinal disorders. In the treatment of disorders associated with increased MEMAP expression or activity, it is desirable to decrease the expression or activity of MEMAP. In the treatment of disorders associated with decreased MEMAP expression or activity, it is desirable to increase the expression or activity of MEMAP.

Therefore, in one embodiment, MEMAP or a fragment or derivative thereof may be 35 administered to a subject to treat or prevent a disorder associated with decreased expression or

activity of MEMAP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, 5 leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory 10 distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, 15 glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, 20 Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural 25 muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, 30 encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, 35 endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including

mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a gastrointestinal disorder such as dysphagia, peptic esophagitis, 5 esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, 10 ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha₁-antitrypsin deficiency, Reye's syndrome, primary 15 sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas.

In another embodiment, a vector capable of expressing MEMAP or a fragment or derivative 20 thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MEMAP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified MEMAP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MEMAP including, but not limited to, 25 those provided above.

In still another embodiment, an agonist which modulates the activity of MEMAP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MEMAP including, but not limited to, those listed above.

In a further embodiment, an antagonist of MEMAP may be administered to a subject to treat 30 or prevent a disorder associated with increased expression or activity of MEMAP. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, neurological and gastrointestinal disorders described above. In one aspect, an antibody which specifically binds MEMAP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express MEMAP.

35 In an additional embodiment, a vector expressing the complement of the polynucleotide

encoding MEMAP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MEMAP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate 5 therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

10 An antagonist of MEMAP may be produced using methods which are generally known in the art. In particular, purified MEMAP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind MEMAP. Antibodies to MEMAP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, 15 and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with MEMAP or with any fragment or oligopeptide 20 thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to 25 MEMAP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of MEMAP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

30 Monoclonal antibodies to MEMAP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and 35 Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce MEMAP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

10 Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

15 Antibody fragments which contain specific binding sites for MEMAP may also be generated. For example, such fragments include, but are not limited to, F(ab')₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

20 Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between MEMAP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering MEMAP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

30 Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for MEMAP. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of MEMAP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple MEMAP epitopes, represents the average affinity, or avidity, of the antibodies for MEMAP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular MEMAP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in

which the MEMAP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_d ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of MEMAP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical

5 Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, 10 preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of MEMAP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al., supra.)

In another embodiment of the invention, the polynucleotides encoding MEMAP, or any 15 fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding MEMAP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of 20 sequences encoding MEMAP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totowa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense 25 sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other 30 gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding MEMAP may be used for 35 somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency

(e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475),
5 cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated
10 cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the
15 case where a genetic deficiency in MEMAP expression or regulation causes disease, the expression of MEMAP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in MEMAP are treated by constructing mammalian expression vectors encoding MEMAP and
20 introducing these vectors by mechanical means into MEMAP-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of MEMAP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). MEMAP may be expressed
30 using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the
35 ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the

FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding MEMAP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID 5 TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of 10 these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to MEMAP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding MEMAP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive 15 element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for 20 receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and A.D. Miller (1988) *J. Virol.* 62:3802-3806; Dull, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") 25 discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; 30 Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver 35 polynucleotides encoding MEMAP to cells which have one or more genetic abnormalities with respect to the expression of MEMAP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have

proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. 5 (1999) *Annu. Rev. Nutr.* 19:511-544; and Verma, I.M. and N. Somia (1997) *Nature* 389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding MEMAP to target cells which have one or more genetic abnormalities with respect to the expression of MEMAP. The use of herpes simplex virus (HSV)-based vectors may be 10 especially valuable for introducing MEMAP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 69:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. 15 Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 20 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation 25 of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding MEMAP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During 30 alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for MEMAP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of 35 MEMAP-coding RNAs and the synthesis of high levels of MEMAP in vector transduced cells. While

alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will 5 allow the introduction of MEMAP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

10 Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using 15 triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

20 Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze 25 endonucleolytic cleavage of sequences encoding MEMAP.

25 Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for 30 secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA 35 sequences encoding MEMAP. Such DNA sequences may be incorporated into a wide variety of

vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible 5 modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, 10 cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding MEMAP. Compounds which may be effective in altering expression of a specific polynucleotide may include, 15 but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased 20 MEMAP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding MEMAP may be therapeutically useful, and in the treatment of disorders associated with decreased MEMAP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding MEMAP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in 25 altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a 30 library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding MEMAP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding MEMAP are assayed by any method commonly known in the art. Typically, the expression of a 35 specific nucleotide is detected by hybridization with a probe having a nucleotide sequence

complementary to the sequence of the polynucleotide encoding MEMAP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test

5 compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular

10 embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable

15 for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

20 Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable

25 excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of MEMAP, antibodies to MEMAP, and mimetics, agonists, antagonists, or inhibitors of MEMAP.

30 The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form.

35 These compositions are generally aerosolized immediately prior to inhalation by the patient. In the

case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, 5 J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

10 Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising MEMAP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, MEMAP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to 15 transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) *Science* 285:1569-1572).

20 For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

25 A therapeutically effective dose refers to that amount of active ingredient, for example MEMAP or fragments thereof, antibodies of MEMAP, and agonists, antagonists or inhibitors of MEMAP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are 30 used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

35 The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the

active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, 5 or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μ g to 100,000 μ g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their 10 inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind MEMAP may be used for the diagnosis of disorders characterized by expression of MEMAP, or in assays to monitor patients being 15 treated with MEMAP or agonists, antagonists, or inhibitors of MEMAP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for MEMAP include methods which utilize the antibody and a label to detect 20 MEMAP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring MEMAP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of MEMAP expression. Normal or standard values for MEMAP expression are established by combining body 25 fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to MEMAP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of MEMAP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for 30 diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding MEMAP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect 35 and quantify gene expression in biopsied tissues in which expression of MEMAP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess

expression of MEMAP, and to monitor regulation of MEMAP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding MEMAP or closely related molecules may be used to identify nucleic acid sequences which encode MEMAP. The specificity of the probe, whether 5 it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding MEMAP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% 10 sequence identity to any of the MEMAP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:38-74 or from genomic sequences including promoters, enhancers, and introns of the MEMAP gene.

Means for producing specific hybridization probes for DNAs encoding MEMAP include the cloning of polynucleotide sequences encoding MEMAP or MEMAP derivatives into vectors for the 15 production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

20 Polynucleotide sequences encoding MEMAP may be used for the diagnosis of disorders associated with expression of MEMAP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including 25 adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), 30 Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, 35 erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves'

disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura,

5 ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders,

10 progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the

15 nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic,

20 endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a gastrointestinal disorder such as dysphagia, peptic esophagitis,

25 esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis,

30 ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha₁-antitrypsin deficiency, Reye's syndrome, primary

35 sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis,

peliosis hepatitis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas. The polynucleotide sequences encoding MEMAP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered MEMAP expression. Such qualitative or quantitative methods are well known in the art.

5 In a particular aspect, the nucleotide sequences encoding MEMAP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding MEMAP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding MEMAP in the 10 sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to 15 monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of 20 MEMAP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding MEMAP, under conditions suitable for hybridization or 25 amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard 30 values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, 35 hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or 35 overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals

to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding MEMAP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding MEMAP, or a fragment of a polynucleotide complementary to the polynucleotide encoding MEMAP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding MEMAP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding MEMAP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of MEMAP include radiolabeling or biotinylation of nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the

polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. The microarray may also be 5 used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate 10 and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, antibodies specific for MEMAP, or MEMAP or fragments thereof may be used as elements on a microarray. The microarray may be used to monitor or measure 15 protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by 20 quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the 25 hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, 30 as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed 35 molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and

toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The

optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic 5 cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for MEMAP to quantify 10 the levels of MEMAP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendoza, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a 15 thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, 20 N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

25 In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic 30 response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are 35 incubated with antibodies specific to the polypeptides of the present invention. The amount of

protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

5 Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are
10 well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding MEMAP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may
15 be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes
20 (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism
25 (RFLP). (See, e.g., Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding MEMAP on a
30 physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse,
35 may reveal associated markers even if the exact chromosomal locus is not known. This information is

valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further 5 investigation. (See, e.g., Gatti, R.A. et al. (1988) *Nature* 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, MEMAP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug 10 screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between MEMAP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT 15 application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with MEMAP, or fragments thereof, and washed. Bound MEMAP is then detected by methods well known in the art. Purified MEMAP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a 20 solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding MEMAP specifically compete with a test compound for binding MEMAP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with MEMAP.

25 In additional embodiments, the nucleotide sequences which encode MEMAP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding 30 description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/149,641 and U.S. Ser. No. 60/164,203 are hereby expressly incorporated 35 by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Genomics, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1

ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) *Anal. Biochem.* 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 5 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSCAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows.

10 Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (PE Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI 15 PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (PE Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the 20 cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VI.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the 25 art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other 30 parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default 35 parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned

sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried 5 against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA.

10 The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene 15 families. (See, e.g., Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:38-74. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

20 **IV. Analysis of Polynucleotide Expression**

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel, 1995, *supra*, ch. 4 and 16.)

25 Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

30
$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum}\{\text{length(Seq. 1)}, \text{length(Seq. 2)}\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is 35 calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the

product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate 5 the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% 10 identity and 100% overlap.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding MEMAP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, 15 reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

20 V. Chromosomal Mapping of MEMAP Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:38-74 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:38-74 were assembled into clusters of contiguous and overlapping sequences using 25 assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

30 The genetic map locations of SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:60, SEQ ID NO:63, and SEQ ID NO:67 are described in The Invention as ranges, or intervals, of human chromosomes. More than one map location is reported for SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:44, SEQ ID NO:60, and SEQ ID NO:63, indicating that previously mapped sequences having similarity, but not complete identity, to SEQ ID NO:38 ,SEQ 35 ID NO:39, SEQ ID NO:44, SEQ ID NO:60, and SEQ ID NO:63 were assembled into their respective

clusters. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.)

5 The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Diseases associated with the public and Incyte sequences located within the indicated intervals are also reported in the Invention where applicable. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be

10 employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VI. Extension of MEMAP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:38-74 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this

15 fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would

20 result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction

25 mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg²⁺, (NH₄)₂SO₄, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the

30 alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE

35 and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar,

Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

5 The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) 10 agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 15 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; 20 Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems).

25 In like manner, the polynucleotide sequences of SEQ ID NO:38-74 are used to obtain 5' regulatory sequences using the procedure above, along with oligonucleotides designed for such extension, and an appropriate genomic library.

VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:38-74 are employed to screen cDNAs, 30 genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of $[\gamma-^{32}P]$ adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase 35 (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a

SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

5 The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and
10 compared.

VIII. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the
15 aforementioned technologies should be uniform and solid with a non-porous surface (Schena (1999), *supra*). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to
20 those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be
25 selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser
30 desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

35 Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and

poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37 °C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85 °C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

15 **Microarray Preparation**

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR 25 Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 μ l of the array element DNA, at an average 30 concentration of 100 ng/ μ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate 35 buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60 °C followed by washes in

0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65 °C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60 °C. The arrays are washed for 10 min at 45 °C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45 °C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital

(A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

5 A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The 10 software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

IX. Complementary Polynucleotides

Sequences complementary to the MEMAP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring MEMAP. Although use of 15 oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of MEMAP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a 20 complementary oligonucleotide is designed to prevent ribosomal binding to the MEMAP-encoding transcript.

X. Expression of MEMAP

Expression and purification of MEMAP is achieved using bacterial or virus-based expression systems. For expression of MEMAP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA 25 transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express MEMAP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of MEMAP in eukaryotic cells is achieved by infecting 30 insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding MEMAP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to 35 infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases.

Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, MEMAP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from MEMAP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified MEMAP obtained by these methods can be used directly in the assays shown in Examples XI and XV.

XI. Demonstration of MEMAP Activity

MEMAP activity is demonstrated using a generic immunoblotting strategy or through a MEMAP-specific activity assay as outlined below. As a general approach, cell lines or tissues transformed with a vector containing MEMAP coding sequences can be assayed for MEMAP activity by immunoblotting. Transformed cells are denatured in SDS in the presence of β -mercaptoethanol, nucleic acids are removed by ethanol precipitation, and proteins are purified by acetone precipitation. Pellets are resuspended in 20 mM tris buffer at pH 7.5 and incubated with Protein G-Sepharose precoated with an antibody specific for MEMAP. After washing, the Sepharose beads are boiled in electrophoresis sample buffer, and the eluted proteins subjected to SDS-PAGE. Proteins are transferred from the SDS-PAGE gel to a membrane for immunoblotting, and the MEMAP activity is assessed by visualizing and quantifying bands on the blot using antibody specific for MEMAP as the primary antibody and 125 I-labeled IgG specific for the primary antibody as the secondary antibody.

A specific assay for MEMAP activity measures the expression of MEMAP on the cell surface. cDNA encoding MEMAP is transfected into a mammalian (non-human) cell line. Cell surface proteins are labeled with biotin as described in de la Fuente, M.A.. et al. ((1997) Blood 90:2398-2405). Immunoprecipitations are performed using MEMAP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of MEMAP expressed on the cell surface.

35 In an alternative specific assay, MEMAP transport activity is assayed by measuring uptake of

labeled substrates into Xenopus laevis oocytes. Oocytes at stages V and VI are injected with MEMAP mRNA (10 ng per oocyte) and incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM CaCl₂, 1mM MgCl₂, 1mM Na₂HPO₄, 5 mM Hepes, 3.8 mM NaOH , 50µg/ml gentamycin, pH 7.8) to allow expression of MEMAP protein. Oocytes are then transferred 5 to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM CaCl₂, 1mM MgCl₂, 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., amino acids, sugars, drugs, and neurotransmitters) is initiated by adding a ³H substrate to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the oocytes three times in Na⁺-free medium, measuring the incorporated ³H, and comparing with controls. MEMAP activity is proportional to the level of internalized ³H substrate.

10

XII. Functional Assays

MEMAP function is assessed by expressing the sequences encoding MEMAP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA 15 expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 µg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 µg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein 20 provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or 25 CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as 30 measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of MEMAP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding MEMAP and either CD64 or CD64-GFP. 35 CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions

of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding MEMAP and other genes of interest can be analyzed by 5 northern analysis or microarray techniques.

XIII. Production of MEMAP Specific Antibodies

MEMAP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

10 Alternatively, the MEMAP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

15 Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (PE Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti- 20 MEMAP activity by, for example, binding the peptide or MEMAP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIV. Purification of Naturally Occurring MEMAP Using Specific Antibodies

Naturally occurring or recombinant MEMAP is substantially purified by immunoaffinity chromatography using antibodies specific for MEMAP. An immunoaffinity column is constructed by 25 covalently coupling anti-MEMAP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing MEMAP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of MEMAP (e.g., high ionic strength 30 buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/MEMAP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and MEMAP is collected.

XV. Identification of Molecules Which Interact with MEMAP

MEMAP, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter 35 reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate

molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled MEMAP, washed, and any wells with labeled MEMAP complex are assayed. Data obtained using different concentrations of MEMAP are used to calculate values for the number, affinity, and association of MEMAP with the candidate molecules.

5 Alternatively, molecules interacting with MEMAP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, *Nature* 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

 MEMAP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions
10 between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

 Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the
15 invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	38	112301	PITUNOT01	003382R1 (HMC1NOT01), 094523R1 (PITUNOT01), 112301H1 (PITUNOT01), 301778X11 (TESTNOT04), 320551X13 (EOSIHE02), 1368852R1 (SCORNON02), 1800784H1 (COLNNOT27), 2117174X17C1 (BRSTTUT02), 25226345F6 (BRAITUT21), 4333609H1 (KIDCTMT01)
2	39	997947	KIDNTUT01	997947H1 (KIDNTUT01), 997947T6 (KIDNTUT01), 1417936X306D1 (KIDNNOT09), 1672062X307V1 (BLADNOT05), 3738956T6 (MENTNOT01), SCCA01437V1, SCCA05013V1, SCCA01691V1, SCCA02873V1
3	40	1521513	BLADTUT04	1222062H1 (NEUTGM01), 1521513H1 (BLADTUT04), 1521513T1 (BLADTUT04), 3558522F6 (LUNGNOT31), 3558522T6 (LUNGNOT31)
4	41	1863994	PROSNOT19	265171R6 (HNT2AGT01), 1863994H1 (PROSNOT19), 3750444F6 (UTRNOT18), 4177677F6 (BRAINNOT22), 4697638F6 (BRAINNOT01), 4774040F6 (BRAQNOT01), SCEA02960V1
5	42	2071941	ISLTNOT01	286350R1 (EOSIHE02), 491305R1 (HNT2AGT01), 724168R1 (SYNOOAT01), 1466668F1 (PANCUT02), 2071941H1 (ISLTNOT01), 2071941X11C1 (ISLTNOT01), 3579445H1 (293TF3T01)
6	43	2172512	ENDCNOT03	2172512H1 (ENDCNOT03), 2544419F6 (UTRSNOT11), 2798626H1 (NPOLNOT01), 3203359H1 (PENCNOT02), 91241299
7	44	2483172	SMCANOT01	217987F1 (STOMNOT01), 1289703F6 (BRAINNOT11), 1289703T6 (BRAINNOT11), 2211377F6 (SINTFET03), 2483172H1 (SMCANOT01), 2493236H1 (ADRETUT05), 3274006F6 (PROSBPT06)

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
8	45	2656128	THYMNOT04	2654722T6 (THYMNOT04), 2656128H1 (THYMNOT04), 2837168F6 (TLYMMNOT03)
9	46	5855841	FIBAUNNOT02	894553T1 (BRSTNOT05), 1296289F1 (PGANNNOT03), 1466541T1 (PANCUT02), 2046927F6 (THP1T7T01), 2058873R6 (OVARNOT03), 3800875F6 (SPLNNNOT12), 5855841H1 (FIBAUNNOT02)
10	47	603462	BRSTTUT01	603462H1 (BRSTTUT01), 1487733H1 (UCMCL5T01), 1750451F6 (STOMTUT02), 5182853T6 (LUNGTM03)
11	48	747681	BRAITUT01	747681H1 (BRAITUT01), 752009R1 (BRAITUT01), 1267874F1 (BRAINNOT09), 1833308R6 (BRAINON01), 2673538X19F1 (KIDNNNOT19), SBCA07003F3, SCDA07521V1, SCDA04982V1, SCDA07275V1
12	49	919469	RATRNOT02	153337R6 (THP1PLB02), 1525415F6 (UCMCL5T01), 1527804F1 (UCMCL5T01), 1985565R6 (LUNGAST01), 2397811T6 (THP1AZT01), SARB01416F1, SARA03198F1
13	50	977658	BRSTNOT02	977658H1 (BRSTNOT02), 1873689F6 (LEUKNOT02), 2155095F6 (BRAINNOT09), 2186432F6 (PROSNOT26), 2204117F6 (SPLNFET02), 2206291F6 (SPLNFET02), 325504BR6 (OVARTUN01), 3501520H1 (ADRENNOT11), 3743427H1 (THYMMNOT08)
14	51	1004703	BRSTNOT03	742178H1 (PANCNOT04), 1444583F6 (THYRNNOT03), 2068902X15C1 (ISLTNOT01), 2616367F6 (GBLANOT01), SBVA02190V1
15	52	1334051	COLNNOT13	3222815T6 (COLNNON03), SXBC00794V1, SXBC00639V1
16	53	1336728	COLNNOT13	630458R6 (KIDNNNOT05), 1336728H1 (COLNNOT13), SXBC00758V1, SXBC01825V1, SXBC01531V1, SXBC01624V1, SXBC00128V1

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
17	54	1452856	PENITUT01	873008R1 (LUNGCAST01), 1452856H1 (PENITUT01), 2433573H1 (BRAVUNT02), 2444932F6 (THP1NOT03), 2858295F6 (SININOT03)
18	55	1562471	SPANNNOT04	286237F1 (ECOSIHET02), 1562471H1 (SPLNNNOT04), 1880730F6 (LEUKNOT03), 3420608F6 (UCMCNOT04), SBWA00968V1, SXBC01387V1, SBWA02301V1
19	56	1618158	BRAITUT12	967563R1 (BRSTNOT05), 1618158H1 (BRAITUT12), 1785271F6 (BRAINNOT10), 2074680F6 (ISLTNOT01), 2822196H1 (ADRETUT06)
20	57	1656935	URETTUT01	1656935F6 (URETTUT01), 1656935H1 (URETTUT01), 2827605F6 (TLYMNNOT03), 5272146H1 (OVARDIN02), 91482116
21	58	1859305	PROSNOT18	079372F1 (SYNORAB01), 639845R1 (BRSTNOT03), 1859305H1 (PROSNOT18), 3328091F6 (HEAONOT04), 3354812F6 (PROSNOT28), 5510642H1 (BRADDIR01)
22	59	1949083	PITUNOT01	1287161H1 (BRAINNOT11), 1949083H1 (PITUNOT01), 1949083R6 (PITUNOT01), 1949083T6 (PITUNOT01), 3814131F6 (TONSNOT03)
23	60	1996357	BRSTTUT03	260527R6 (HNT2RAT01), 260527T6 (HNT2RAT01), 1313441F1 (BLADTUT02), 1442781R1 (THYRNOT03), 1996357H1 (BRSTTUT03), 1996357T6 (BRSTTUT03), 4262451H1 (BSCNDIT02), SAZA00147F1
24	61	2061330	OVARNOT03	2061330H1 (OVARNOT03), 2724233T6 (LUNGUT10), 5104031T6 (PROSTUS20)
25	62	2346947	TESTTUT02	2346947F6 (TESTTUT02), 2346947H1 (TESTTUT02), 4051345F6 (SINTNOT18)

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
26	63	2795577	NPOLNOT01	867213R6 (BRAITUT03), 2381770H1 (ISLTNOT01), 2795577CT1 (NPOLNOT01), 2795577H1 (NPOLNOT01)
27	64	3255825	OVARTUN01	3255825CT1 (OVARTUN01), 3255825H1 (OVARTUN01)
28	65	3393430	LUNGNOT28	2187169H1 (PROSNOT26), 3393256H1 (LUNGNOT28), 3393430H1 (LUNGNOT28), 3395774H1 (LUNGNOT28), 4689688H1 (LIVRTUT12), 4895996H1 (LIVRTUT12), 4896461F6 (LIVRTUT12), 4984527F6 (LIVRTUT10), 4992946H1 (LIVRTUT11)
29	66	3490990	EPIGNOT01	1235428F1 (LUNGFET03), 1662973T6 (BRSTNOT09), 2362021H1 (LUNGFET05), 2362021R6 (LUNGFET05), 3490990H1 (EPIGNOT01)
30	67	3635154	LIVRNOT03	027592H1 (SPLNFET01), 3635154H1 (LIVRNOT03), 91012932
31	68	4374347	CONFNOT03	860875X11 (BRAITUT03), 898143R6 (BRSTNOT05), 4374347H1 (CONFNOT03)
32	69	4596747	COLSTUT01	137213R1 (SYNORAB01), 545568R6 (OVARNOT02), 1235402F1 (LUNGFET03), 1268010F1 (BRAINNOT09), 1271078F1 (TESTTUT02), 1301951F6 (BRSTNOT07), 1994442R6 (BRSTTUT03), 2343102H1 (TESTTUT02), 3274538F6 (PROSBPT06), 4596747H1 (COLSTUT01)
33	70	5052680	BRSTNOT33	1973688H1 (UCMCL5T01), 3926410F6 (KIDNNNOT19), 4501839F6 (BRAVITXT02), 5052680F6 (BRSTNOT33), 5052680H1 (BRSTNOT33), 5186780F6 (LUNGMT04)
34	71	5373575	BRAINOT22	262776T6 (HNT2AGT01), 1234057F1 (LUNGFET03), 1741526R6 (HIFONON01), 1871204F6 (SKINBIT01), 2192479F6 (THYRTUT03), 2556849H1 (THYMNNOT03), 2722451T6 (LUNGUT10), 4114985H1 (UTRSTUT07), 5373575H1 (BRAINOT22)

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
35	72	5524468	LIVRDIRO1	4024068F6 (BRAXNOT02), 5524468H1 (LIVRDIR01), SXBC01952V1
36	73	5944279	COLADIT05	1662182H1 (BRSTNOT09), 1698677F6 (BLADTUT05), 1916639R6 (PROSNOT06), 1916639T6 (PROSNOT06), 2298565R6 (BRSTNOT05), 2298565T6 (BRSTNOT05), 2583019F6 (KIDNTUT13), 2870903F6 (THYRNOT10), 3970715H1 (PROSTUT10), 3971695H1 (PROSTUT10), 5944279H1 (COLADIT05)
37	74	6114480	SINITMT04	1579843F6 (DUODNOT01), 1579843T6 (DUODNOT01), 4181024T6 (SINITTUT03), 6114480H1 (SINITMT04), SXBC00007V1, SXBC00504V1, SCSA05104V1

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
1	351	S31 T116 S169 T229 T2 S209 T306	N128	Signal peptide: M1-A33	Paraneoplastic neuronal antigen MA1 [Homo sapiens] g4104634	BLAST-GenBank MOTIFS SPSCAN
2	458	T198 S27 S37 T87 S251 S257 S325 S373 S405 S422 T454 T210 S228 S401 Y93	N75 N159 N279 N445	Signal peptide: M1-T24 Glycoprotein signature: C199-L448	Pancortin-3 [Mus musculus] 93218528	BLAST-GenBank MOTIFS SPSCAN HHMER BLAST-PRODOM
3	219	T51 S120 S163 T175 T181 S3 T12 T45 S75 S104 S128	N2 N62 N107	Signal peptide: M1-C42 Transmembrane domain: L32-F49 C-type lectin domain: C80-E206	Murine macrophage C-type lectin [Mus musculus] g5821286	BLAST-GenBank MOTIFS SPSCAN HHMER HHMER-PFAM BLIMPS-BLOCKS PROFILESCAN BLAST-DOMO
4	276	S213 S91 S113 S35 S70 S76 S147 T163 S206		Signal peptide: M1-G31 Transmembrane domain: I184-F201 Cell attachment sequence: R149-D151		BLAST-GenBank MOTIFS HHMER

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
5	375	S18 S205 T286 S3 S120 S197 T260 Y85		Transmembrane domains: W139-R158; F173-H191 P232-Q254 Transmembrane protein signature: I95-C369	Transmembrane protein [S. pombe] g1065898	BLAST-GenBank MOTIFS HMMER BLAST-DOMO BLAST-PRODOM
6	249	T7 T135 T170 S204 Y154	N18 N92 N147		Phospholipid scramblase [Homo sapiens] g4092081	BLAST-GenBank MOTIFS
7	353	T162 T4 S97 T115 S165 S194 T225 S242 S17 S47 S205	N299	Signal peptide: M1-A33	Paraneoplastic neuronal antigen MA1 [Homo sapiens] g4104634	BLAST-GenBank MOTIFS SPSCAN
8	194	T12 S115 S29 S99 S187	N95 N147	Signal peptide: M1-C50 Transmembrane domain: L38-L56 C-type lectin domain: C75-E194	Lectin-like NK cell receptor LILT1 [Homo sapiens] g6651065	BLAST-GenBank MOTIFS SPSCAN HMMER BLIMPS-BLOCKS BLAST-DOMO
9	322	S304 S48 S146 S72 T133 S255 S280	N20 N60 N70	Signal peptide: M1-A50		BLAST-GenBank MOTIFS SPSCAN

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
10	335	S125 S140 S183 S222 T252		Transmembrane domains: G71-L94; A255-I283 GufA transmembrane protein domain: L12-H101; G180-G335 Glycosaminoglycan attachment site: S310-G313	GufA protein [Thermotoga maritima] g4982315	BLAST-GenBank MOTIFS HMMER BLAST-PRODOM BLAST-DOMO
11	620	S49 S108 T146 S300 T348 T349 S607 S4 S128 S183 S234 T420 S460 S467 S543 Y597	N144 N202 N264 N274 N293 N341 N492 N505 N526 N542	Transmembrane domain: M563-W582 Immunoglobulin domain: G439-A499 Leucine-rich repeat signature: L337-L350 Glycoprotein hormone receptor domain: T40-L198	Slit2 [Rattus norvegicus] g4585574	BLAST-GenBank MOTIFS HMMER HMMER-PFAM BLIMPS-PRINTS BLAST-DOMO

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
12	491	T231 T232 S253 T482 S185 S276	N56 N220 N229	Transmembrane domains: I115-I142; I184-V201 F422-F441 Transmembrane protein domain: L8-Y215; I396-F471	Selectively expressed in embryonic epithelia protein-1 [Mus musculus] g6715148	BLAST-GenBank MOTIFS HMMER BLAST-PRODOM
13	580	S557 S10 T34 S51 T92 T210 S343 T12 S217 T222 S268 S296 T417 T523 S550	N159 N179 N220 N230	Transmembrane domains: F297-F313; I356-I373 L496-I514 Lipases serine active site: L104-A113	PB39 [Homo sapiens] g3462515	MOTIFS HMMER
14	455	T53 T182 S239 S69 S135 S202 T280 S355 S372 Y38	N67 N180 N243	Transmembrane domains: V81-V99; I343-I361 S375-V392; W425-Y442 Glycosaminoglycan attachment site: S162-G165	putative G-protein coupled receptor [Homo sapiens] g6649579	BLAST-GenBank MOTIFS HMMER

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
15	277	S265 T66 T225 S268 S273 S30 S49 S61 S152 S193 Y242	N29 N38 N47 N48 N92 N160 N210	Transmembrane domain: K9-F27 Brush border protein domain: Y8-R277 RGD cell attachment sequence: R113-D115	AdRab-A brush border membrane protein [Oryctolagus cuniculus] g1762	BLAST-GenBank MOTIFS HMMER BLAST-PRODOM
16	647	S490 T50 S67 S105 T110 S121 T220 S249 S264 S272 S322 T389 S469 T501 S639 S132 T155 S242 S324 T381 T400 S522 S554	N261	Signal peptide: M1-A22 Transmembrane domains: L328-L347; M406-L424 L559-A578; W618-L638 GufA transmembrane protein domain: E485-L640 Glycossaminoglycan attachment site: S34-G37	LIV-1 protein [Homo sapiens] g1256001	BLAST-GenBank MOTIFS SPSCAN HMMER BLAST-PRODOM
17	406	S29 S215 S236 T69	N23	Transmembrane domains: Q76-V95; W286-S313 M367-I384		MOTIFS HMMER

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
18	290	T221 S44 S69 S71 S81 T94 T101 T113 T131 S216 Y284	N88	Signal peptide: M1-A19 Transmembrane domains: P160-M181 Immunoglobulin domain: R33-I110 Transmembrane glycoprotein domain: I22-D116	NK inhibitory receptor [Homo sapiens] g6707799 CMRF-35-H9 leukocyte antigen [Homo sapiens] g4103066	BLAST-GenBank MOTIFS SPSCAN HMMER HMMER-PFAM BLAST-PRODOM BLAST-DOMO
19	390	S7 T68 S153 T23 T166 T281 Y20 Y37	N5 N88 N330 N367	Immunoglobulins and MHC proteins signature: T90-P112; F242-V259 Glycoprotein antigen signature: L61-V72; V92-I113		MOTIFS BLIMPS-BLOCKS BLIMPS-PRODOM
20	427	S13 S41 S65 S66 S99 T150 S323 S324 S101 S275 S353 S367 T399 Y71	N106 N148 N171 N233 N312	Mucin glycoprotein precursor domain: V136-P142	Gastric mucin [Sus scrofa] g915208	BLAST-GenBank MOTIFS BLIMPS-PRODOM

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
21 459	T4 S60 S66 S116 T176 S16 T235	N14 N158 N323	Transmembrane domains: F202-V219; I246-L268 W343-L367; P417-P440	six transmembrane epithelial antigen of prostate [Homo sapiens] g6572948	BLAST-GenBank MOTIFS HMMER	
22 229	S13 S118 T155 Y24		Transmembrane domains: I93-V111; V132-L150 F164-V182	Transmembrane protein domain: S156-V182	MOTIFS HMMER BLIMPS-PRODOM	
23 311	S85 S234 S236 S269 S80 S119 S186 T294	N22	Transmembrane domains: W58-I176; P152-K177 A216-Y232		MOTIFS HMMER	
24 92	S47 T54 T12 S70	N62		HERV-E envelope glycoprotein [Homo sapiens] g2587024	BLAST-GenBank MOTIFS	
25 258	S34 T33 S148 S243		Transmembrane domains: I39-I57; F86-L106 V122-I140; L190-S210		MOTIFS HMMER	

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosyla- tion Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
26	226	S56 S128 T196 T167 Y194	N54 N187 N198	Signal peptide: M1-P50 Transmembrane domains: T23-L43; M72-A89 I101-I124; I158-N178 Transmembrane 4 family signature: A70-I120 Lysosomal-associated transmembrane protein domain: C15-Y223	MTP (mouse transporter protein) [Mus musculus] g1276631	BLAST-GenBank MOTIFS SPSCAN HMMER PROFILESCAN BLAST-PRODOM
27	136	S3 S132		Signal peptide: M1-R53 Transmembrane domains: I10-L28; T26-I50 F70-L89 Transmembrane protein domain: D31-Y104		MOTIFS SPSCAN HMMER BLAST-PRODOM

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
28	458	T408 T98 S126 S170 T334	N96 N151 N293 N332	Signal peptide: M1-A20 Transmembrane domain: L10-N30 Membrane glycoprotein signature: L9-V101; L64-Q457 Olfactory ligand binding domain: T67-S452	Potential ligand (odorant) binding protein [Rattus rattus] g57732	BLAST-GenBank MOTIFS SPSCAN HHMER BLAST-PRODOM BLAST-DOMO
29	368	S24 T166 T302 S12 S134 Y307	N17		Fuzzy (TM protein involved in tissue polarity) (Drosophila melanogaster) g2564657	BLAST-GenBank MOTIFS
30	91	T44 S84		Signal peptide: M1-A19 Transmembrane domain: P58-S82 Glycophorin A proteins signature: T22-S32; I63-G91 Glycophorin domain: M1-R86	Preglycophorin B [Homo sapiens] g4803699	BLAST-GenBank MOTIFS SPSCAN HHMER BLIMPS-BLOCKS PROFILESCAN BLAST-PRODOM BLAST-DOMO

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
31	295	S96 T113 S129 T155 T125 T157 T187 S222 T231 T263 Y212	N111 N169 N223	Signal peptide: M1-G48 Transmembrane domain: L241-L259 Immunoglobulin domain: K159-V216 Carcinoembryonic antigen domain: I38-P147 Glycoprotein antigen domain: M1-V140; Y141-Y234 G239-S295	Biliary glycoprotein [Mus musculus] g312590	BLAST-GenBank MOTIFS SPSCAN HHMER HHMER-PFAM BLAST-PRODOM BLAST-DOMO
32	724		T39 S47 T171 S205 T224 S225 T241 S285 S301 T323 S352 T353 S439 S509 S517 S537 T659 T707 S8 S18 S49 S72 T85 T159 S173 S271 S367 S560 S588 Y499	N279 N348 Transmembrane domain: I611-F630 Membrane protein domain: T4-L209	MOTIFS HHMER BLAST-DOMO	

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
33	331	S117 S147 S149 T320 S138 S174 T274 T319 S328 Y198	N222	Signal peptide: M1-S16 Transmembrane domains: A67-N87; I118-C134 W240-V269; L294-Y310 Transmembrane protein domain: A6-T311	Putative Golgi UDP-GlcNAc transporter [<i>S. pombe</i>] g3738167	BLAST-GenBank MOTIFS SPSCAN HMMER BLAST-PRODOM
34	398	T42 T158 S271 S28 S285 T334 S375		Transmembrane domain: I59-L79 Band 7 family domain: F64-A231, A78-V90; R116-L154 Stomatin signature: T84-L106; L131-P152 T166-L183; I186-G209 L54-Q227	Stomatin-like protein UNC24 [<i>Homo sapiens</i>] g5326747	BLAST-GenBank MOTIFS HMMER HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO
35	220	S199 T120 S192	N107	Signal peptide: M1-G19 Leucine rich repeats: A62-F85; Q86-S109 G110-G133; A134-R157 A158-S181; H184-P207	Similar to Leucine-rich transmembrane proteins [<i>Homo sapiens</i>] g2781386	BLAST-GenBank MOTIFS SPSCAN HMMER HMMER-PFAM BLIMPS-PRINTS

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
36	706	T564 T74 T113 S291 S452 S632 S14 T42 S66 T115 T142 S286 T551 T575 S701	N101	Transmembrane domains: F158-M178; L344-V368 L425-L442; M478-F498 A581-I604; L641-V665 Glycosaminoglycan attachment site: S223-G226	LAK-4p [Homo sapiens] g7209574	BLAST-GenBank MOTIFS HMMER
37	466	T326 S10 T46 T105 S187 S98 T164 T310 S321 Y388	N368	Signal peptide: M1-G23 Transmembrane domain: A236-I255 SPRY domain: A338-S464; E123-S136 E322-W343; V407-F420 Butyrophilin domain: W19-C114	Butyrophilin like receptor [Homo sapiens] g4587209	BLAST-GenBank MOTIFS SPSCAN HMMER BLIMPS-PFAM BLAST-PRODOM BLAST-DOMO

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
38	844-888	Nervous (0.377) Reproductive (0.180) Cardiovascular (0.115) Gastrointestinal (0.115)	Cancer (0.410) Inflammation/Trauma (0.296) Cell Proliferation (0.131)	PBLUESCRIPT
39	579-623	Developmental (0.400) Musculoskeletal (0.200) Nervous (0.200)	Cancer (0.400) Cell Proliferation (0.400)	PSPORT1
40	336-380	Cardiovascular (0.267) Hematopoietic/Immune (0.200) Endocrine (0.133) Reproductive (0.133)	Cancer (0.400) Inflammation/Trauma (0.400) Cell Proliferation (0.133)	pINCY
41	596-640	Nervous (0.588) Gastrointestinal (0.118) Reproductive (0.118)	Inflammation/Trauma (0.470) Cancer (0.235) Cell Proliferation (0.176)	pINCY
42	1281-1325	Reproductive (0.237) Hematopoietic/Immune (0.145) Nervous (0.145)	Cancer (0.441) Inflammation/Trauma (0.323) Cell Proliferation (0.178)	pINCY
43	227-271	Reproductive (0.444) Dermatologic (0.222) Endocrine (0.111) Gastrointestinal (0.111) Nervous (0.111)	Cancer (0.333) Cell Proliferation (0.222) Inflammation/Trauma (0.222)	pINCY

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
44	1368-1412	Nervous (0.339) Reproductive (0.278) Gastrointestinal (0.104)	Cancer (0.478) Inflammation/Trauma (0.278) Cell Proliferation (0.165)	PINCY
45	543-587	Hematopoietic/Immune (0.500) Gastrointestinal (0.188)	Inflammation/Trauma (0.500) Cancer (0.250) Cell Proliferation (0.188)	PINCY
46	280-324	Reproductive (0.267) Nervous (0.233) Gastrointestinal (0.112)	Cancer (0.483) Inflammation/Trauma (0.345) Cell Proliferation (0.155)	PINCY
47	380-424 875-919	Reproductive (0.412) Gastrointestinal (0.176) Cardiovascular (0.118)	Cancer (0.647) Inflammation/Trauma (0.178)	PSPORT1
48	272-316 1514-1558	Nervous (0.645) Developmental (0.129)	Cancer (0.355) Cell Proliferation (0.258) Neurological (0.194)	PSPORT1
49	282-326 768-812	Hematopoietic/Immune (0.238) Gastrointestinal (0.155) Reproductive (0.143)	Cancer (0.381) Inflammation/Trauma (0.381) Cell Proliferation (0.202)	PSPORT1
50	597-641 1074-1118	Reproductive (0.214) Nervous (0.196) Hematopoietic/Immune (0.143)	Cancer (0.464) Inflammation/Trauma (0.304) Cell Proliferation (0.196)	PSPORT1
51	973-1017	Reproductive (0.266) Nervous (0.234) Hematopoietic/Immune (0.125)	Cancer (0.516) Inflammation/Trauma (0.359) Cell Proliferation (0.109)	PSPORT1

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
52	299-343	Gastrointestinal (1.000)	Cancer (0.500) Inflammation/Trauma (0.500)	PINCY
53	380-424 1199-1243	Gastrointestinal (0.289) Reproductive (0.244) Cardiovascular (0.111) Hematopoietic/Immune (0.111)	Cancer (0.578) Inflammation/Trauma (0.311) Cell Proliferation (0.178)	PINCY
54	1135-1179	Nervous (0.195) Reproductive (0.186) Gastrointestinal (0.144)	Cancer (0.449) Inflammation/Trauma (0.305) Cell Proliferation (0.144)	PINCY
55	325-369 820-864	Hematopoietic/Immune (0.750)	Inflammation/Trauma (0.625) Cancer (0.125)	PINCY
56	487-531 1090-1134	Nervous (0.583)	Cancer (0.458) Inflammation/Trauma (0.250)	PINCY
57	569-613 1360-1405	Reproductive (0.429) Hematopoietic/Immune (0.286) Musculoskeletal (0.143) Urologic (0.143)	Cancer (0.571) Inflammation/Trauma (0.286) Cell Proliferation (0.143)	PINCY
58	272-472 551-595 812-1012 1523-1567	Reproductive (0.350) Nervous (0.150) Cardiovascular (0.100) Gastrointestinal (0.100) Hematopoietic/Immune (0.100) Urologic (0.100)	Cancer (0.500) Inflammation/Trauma (0.500)	PINCY

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
59	217-261	Nervous (0.286) Developmental (0.143) Gastrointestinal (0.143) Hematopoietic/Immune (0.143) Reproductive (0.143)	Inflammation/Trauma (0.428) Cancer (0.357) Cell Proliferation (0.143)	PBLUESCRIPT
60	444-488	Nervous (0.207) Reproductive (0.207) Gastrointestinal (0.130) Hematopoietic/Immune (0.130)	Cancer (0.467) Inflammation/Trauma (0.359) Cell Proliferation (0.163)	PSPORT1
61	643-687	Reproductive (0.464) Endocrine (0.143) Cardiovascular (0.107) Gastrointestinal (0.107)	Cancer (0.500) Inflammation/Trauma (0.321)	PSPORT1
62	146-344 390-434 506-704 786-830	Gastrointestinal (0.500) Hematopoietic/Immune (0.250) Reproductive (0.250)	Cancer (0.750) Inflammation/Trauma (0.250)	PINCY
63	163-207	Reproductive (0.315) Gastrointestinal (0.161) Cardiovascular (0.147)	Cancer (0.594) Cell Proliferation (0.231) Inflammation/Trauma (0.210)	PINCY
64	201-506 525-569 606-912 975-1280 1362-1406	Gastrointestinal (0.455) Cardiovascular (0.273) Reproductive (0.189)	Cancer (0.455) Inflammation/Trauma (0.367) Cell Proliferation (0.189)	PSPORT1

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
65	703-747	Gastrointestinal (0.667) Cardiovascular (0.167) Reproductive (0.167)	Cancer (1.000)	pINCY
66	271-315 319-363	Nervous (0.314) Reproductive (0.314) Developmental (0.114) Urologic (0.114)	Cancer (0.429) Cell Proliferation (0.171) Inflammation/Trauma (0.143)	pINCY
67	319-363	Developmental (0.364) Hematopoietic/Immune (0.364) Gastrointestinal (0.182)	Cell Proliferation (0.727) Cancer (0.273) Inflammation/Trauma (0.182)	pINCY
68	812-856	Reproductive (0.444) Nervous (0.222) Endocrine (0.111) Hematopoietic/Immune (0.111) Musculoskeletal (0.111)	Cancer (0.556) Inflammation/Trauma (0.333)	pINCY
69	596-640 1577-1621	Reproductive (0.255) Nervous (0.184) Developmental (0.122) Gastrointestinal (0.122)	Cancer (0.429) Inflammation/Trauma (0.337) Cell Proliferation (0.255)	pINCY
70	379-675 703-747 766-1062 1081-1347	Nervous (0.467) Hematopoietic/Immune (0.200) Reproductive (0.133) Urologic (0.133)	Cancer (0.467) Cell Proliferation (0.267) Inflammation/Trauma (0.267)	pINCY

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
71	18-62	Nervous (0.265) Reproductive (0.206) Musculoskeletal (0.147)	Cancer (0.500) Inflammation/Trauma (0.264) Cell Proliferation (0.147)	PINCY
72	290-488 507-704 759-803	Gastrointestinal (0.333) Hematopoietic/Immune (0.333) Nervous (0.333)	Inflammation/Trauma (0.667) Cancer (0.333)	PINCY
73	649-693 1711-1755	Reproductive (0.392) Gastrointestinal (0.294) Cardiovascular (0.118)	Cancer (0.686) Inflammation/Trauma (0.294)	PINCY
74	704-748	Gastrointestinal (0.923)	Cancer (0.462) Inflammation/Trauma (0.385)	PINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
38	PITUNOR01	This library was constructed using RNA obtained from Clontech (CLON 6584-2, lot 35278). The RNA was isolated from pituitary glands removed from a pool of 18 male and female Caucasian donors, 16 to 70 years old, who died from trauma.
39	KIDNTUT01	This library was constructed using RNA isolated from kidney tumor tissue removed from an 8-month-old female during nephroureterectomy. Pathology indicated Wilms' tumor (nephroblastoma), which involved 90 percent of the renal parenchyma. Prior to surgery, the patient was receiving heparin anticoagulant therapy.
40	BLADTUT04	This library was constructed using RNA isolated from bladder tumor tissue removed from a 60-year-old Caucasian male during a radical cystectomy, prostatectomy, and vasectomy. Pathology indicated grade 3 transitional cell carcinoma in the left bladder wall. Carcinoma in-situ was identified in the dome and trigone. Patient history included tobacco use. Family history included type I diabetes, malignant neoplasm of the stomach, atherosclerotic coronary artery disease, and acute myocardial infarction.
41	PROSNOT19	This library was constructed using RNA isolated from diseased prostate tissue removed from a 59-year-old Caucasian male during a radical prostatectomy with regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 3+3). The patient presented with elevated prostate-specific antigen (PSA). Patient history included colon diverticuli, asbestosis, and thrombophlebitis. Family history included benign hypertension, multiple myeloma, hyperlipidemia and rheumatoid arthritis.
42	ISLTNOT01	This library was constructed using RNA isolated from a pooled collection of pancreatic islet cells.

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
43	ENDCN0R03	This library was constructed using RNA isolated from dermal microvascular endothelial cells removed from a neonatal Caucasian male.
44	SMCAN0R01	This library was constructed using RNA isolated from an aortic smooth muscle cell line derived from the explanted heart of a male obtained during a heart transplant.
45	THYMNOT04	This library was constructed using RNA isolated from thymus tissue removed from a 3-year-old Caucasian male, who died from anoxia.
46	FIBAUNT02	This library was constructed using RNA isolated from untreated aortic adventitial fibroblasts removed from a 65-year-old Caucasian female.
47	BRSTTUT01	This library was constructed using RNA isolated from breast tumor tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma. Patient history included atrial tachycardia and a benign breast neoplasm. Family history included cardiovascular and cerebrovascular disease and depressive disorder.
48	BRAITT01	This library was constructed using RNA isolated from brain tumor tissue removed from a 50-year-old Caucasian female during a frontal lobectomy. Pathology indicated recurrent grade 3 oligoastrocytoma with focal necrosis and extensive calcification. Patient history included a speech disturbance and epilepsy. The patient's brain had also been irradiated with a total dose of 5,082 cgy (Fraction 8). Family history included a brain tumor.
49	RATRN0T02	This library was constructed using RNA isolated from the right atrium tissue of a 39-year-old Caucasian male, who died from a gunshot wound.

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
50	BRSTNOT02	This library was constructed using RNA isolated from diseased breast tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated proliferative fibrocystic changes characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Pathology for the associated tumor tissue indicated an invasive grade 4 mammary adenocarcinoma. Patient history included atrial tachycardia and a benign neoplasm. Family history included cardiovascular and cerebrovascular disease.
51	BRSTNOT03	This library was constructed using RNA isolated from diseased breast tissue removed from a 54-year-old Caucasian female during a bilateral radical mastectomy. Pathology for the associated tumor tissue indicated residual invasive grade 3 mammary ductal adenocarcinoma. Patient history included kidney infection and condyloma acuminatum. Family history included benign hypertension, hyperlipidemia and a malignant neoplasm of the colon.
52	COLNNOT13	This library was constructed using RNA isolated from ascending colon tissue of a 28-year-old Caucasian male with moderate chronic ulcerative colitis.
53	COLNNOT13	This library was constructed using RNA isolated from ascending colon tissue of a 28-year-old Caucasian male with moderate chronic ulcerative colitis.
54	PENITUT01	This library was constructed using RNA isolated from tumor tissue removed from the penis of a 64-year-old Caucasian male during penile amputation. Pathology indicated a fungating invasive grade 4 squamous cell carcinoma involving the inner wall of the foreskin and extending onto the glans penis. Patient history included benign neoplasm of the large bowel, atherosclerotic coronary artery disease, angina pectoris, gout, and obesity. Family history included malignant pharyngeal neoplasm, chronic lymphocytic leukemia, and chronic liver disease.

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
55	SPLNNNOT04	This library was constructed using RNA isolated from the spleen tissue of a 2-year-old Hispanic male, who died from cerebral anoxia.
56	BRAITUT12	This library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated grade 4 gemistocytic astrocytoma.
57	URETTUT01	This library was constructed using RNA isolated from right ureter tumor tissue of a 69-year-old Caucasian male during ureterectomy and lymph node excision. Pathology indicated invasive grade 3 transitional cell carcinoma. Patient history included benign colon neoplasm, tobacco use, asthma, emphysema, acute duodenal ulcer, and hyperplasia of the prostate. Family history included atherosclerotic coronary artery disease, congestive heart failure, and malignant lung neoplasm.
58	PROSNOR18	This library was constructed using RNA isolated from diseased prostate tissue removed from a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrostomy. Pathology indicated adenofibromatous hyperplasia; this tissue was associated with a grade 3 transitional cell carcinoma. Patient history included angina and emphysema. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
59	PITUNNOT01	This library was constructed using RNA obtained from Clontech (CLON 6584-2, lot 35278). The RNA was isolated from the pituitary glands removed from a pool of 18 male and female Caucasian donors, 16 to 70 years old, who died from trauma.

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
60	BRSTTUT03	This library was constructed using RNA isolated from breast tumor tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated multicentric invasive grade 4 lobular carcinoma. The mass was identified in the upper outer quadrant, and three separate nodules were found in the lower outer quadrant of the left breast. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular disease, coronary artery aneurysm, breast cancer, prostate cancer, atherosclerotic coronary artery disease, and type I diabetes.
61	OVARNOT03	This library was constructed using RNA isolated from ovarian tissue removed from a 43-year-old Caucasian female during removal of the fallopian tubes and ovaries. Pathology for the associated tumor tissue indicated grade 2 mucinous cystadenocarcinoma. Patient history included mitral valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, stress reaction, cerebrovascular disease, breast cancer, and uterine cancer.
62	TESTTUT02	This library was constructed using RNA isolated from testicular tumor tissue removed from a 31-year-old Caucasian male during unilateral orchectomy. Pathology indicated embryonal carcinoma.
63	NPOLNOT01	This library was constructed using RNA isolated from nasal polyp tissue removed from a 78-year-old Caucasian male during a nasal polypectomy. Pathology indicated a nasal polyp and striking eosinophilia. Patient history included asthma and nasal polyps.

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
64	OVARTUN01	This normalized library was constructed from 5.36 million independent clones obtained from an ovarian tumor library. RNA was isolated from tumor tissue removed from the left ovary of a 58-year-old Caucasian female during a total abdominal hysterectomy, removal of a single ovary, and inguinal hernia repair. Pathology indicated a metastatic grade 3 adenocarcinoma of colonic origin, forming a partially cystic and necrotic tumor mass in the left ovary, and a nodule in the left mesovarium. A single intramural leiomyoma was identified in the myometrium. The cervix showed mild chronic cystic cervicitis. Patient history included benign hypertension, follicular ovarian cyst, colon cancer, benign colon neoplasm, and osteoarthritis. Family history included emphysema, myocardial infarction, atherosclerotic coronary artery disease, benign hypertension, hyperlipidemia, and primary tuberculous complex. The normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228) and Bonaldo et al. (Genome Research (1996) 6:791).
65	LUNGNOT28	This library was constructed using RNA isolated from lung tissue removed from a 53-year-old male. Pathology for the associated tumor tissue indicated grade 4 adenocarcinoma.
66	EPIGNOT01	This library was constructed using RNA isolated from epiglottic tissue removed from a 71-year-old male during laryngectomy with right parathyroid biopsy. Pathology for the associated tumor tissue indicated recurrent grade 1 papillary thyroid carcinoma.
67	LIVRNOT03	This library was constructed using RNA isolated from liver tissue removed from a Caucasian male fetus, who died from Patau's syndrome (trisomy 13) at 20 weeks' gestation.

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
68	CONFNOT03	<p>This library was constructed using RNA isolated from mesenteric fat tissue removed from a 71-year-old Caucasian male during a partial colectomy and permanent colostomy. Pathology indicated mesenteric fat tissue associated with diverticulosis and diverticulitis with abscess formation. Approximately 50 diverticula were noted, one of which was perforated and associated with abscess formation in adjacent mesenteric fat. The patient presented with atrial fibrillation. Patient history included viral hepatitis, a hemangioma, and diverticulitis of colon. Family history included extrinsic asthma, atherosclerotic coronary artery disease, and myocardial infarction.</p>
69	COLSTUT01	<p>This library was constructed using RNA isolated from colon tumor tissue removed from the sigmoid colon of a 62-year-old Caucasian male during a sigmoidectomy and permanent colostomy. Pathology indicated invasive grade 2 adenocarcinoma, with invasion through the muscularis. Patient history included hyperlipidemia, cataract disorder and dermatitis. Family history included benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, breast cancer, and prostate cancer.</p>
70	BRSTNOT33	<p>This library was constructed using RNA isolated from right breast tissue removed from a 46-year-old Caucasian female during a unilateral extended simple mastectomy with breast reconstruction. Pathology for the associated tumor tissue indicated invasive grade 3 adenocarcinoma, ductal type, with apocrine features, nuclear grade 3 forming a mass in the outer quadrant. There was greater than 50% intraductal component. Patient history included breast cancer.</p>

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
71	BRAINOT22	<p>This library was constructed using RNA isolated from right temporal lobe tissue removed from a 45-year-old Black male during a brain lobectomy. Pathology for the associated tumor tissue indicated dysembryoplastic neuroepithelial tumor of the right temporal lobe. The right temporal region dura was consistent with calcifying pseudotumor of the neuraxis. Patient history included obesity, meningitis, backache, unspecified sleep apnea, acute stress reaction, acquired knee deformity, and chronic sinusitis. Family history included obesity, benign hypertension, cirrhosis of the liver, obesity, hyperlipidemia, cerebrovascular disease, and type II diabetes.</p>
72	LIVRDIR01	<p>This library was constructed using RNA isolated from diseased liver tissue removed from a 63-year-old Caucasian female during a liver transplant. Patient history included primary biliary cirrhosis. Serology was positive for anti-mitochondrial antibody.</p>
73	COLADIR05	<p>This library was constructed using RNA isolated from diseased ascending colon tissue removed from a 32-year-old Caucasian male during a total intra-abdominal colectomy, abdominal-perineal rectal resection, and temporary ileostomy. Pathology indicated chronic ulcerative colitis extending in a continuous fashion from the mid-portion of the ascending colon distally to the rectum. This was characterized microscopically by crypt abscess formation and inflammation confined to the mucosa and submucosa. The terminal ileum exhibited ileitis and the rectal mucosa showed crypt abscess formation. Patient history included tobacco use. Family history included ulcerative colitis, malignant neoplasm of the breast and acute myocardial infarction.</p>

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
74	SINITM04	<p>Library was constructed using RNA isolated from ileum tissue removed from a 70-year-old Caucasian female during right hemicolectomy, open liver biopsy, flexible sigmoidoscopy, colonoscopy, and permanent colostomy. Pathology indicated a non-tumorous margin of ileum. Pathology for the associated tumor indicated invasive grade 2 adenocarcinoma forming an ulcerated mass, situated 2 cm distal to the ileocecal valve. The tumor invaded through the muscularis propria just into the serosal adipose tissue. One (of 16) regional lymph node was positive for a microfocus of metastatic adenocarcinoma. Patient history included a malignant breast neoplasm, type II diabetes, hyperlipidemia, viral hepatitis, an unspecified thyroid disorder, osteoarthritis, and a malignant skin neoplasm. Family history included breast cancer, atherosclerotic coronary artery disease, benign hypertension, cerebrovascular disease, ovarian cancer, and hyperlipidemia.</p>

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	PE Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	PE Biosystems, Foster City, CA;	
ABI AutoAssembler	A program that assembles nucleic acid sequences.	PE Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25:3389-3402.	<i>ESTs</i> : Probability value= 1.0E-8 or less <i>Full Length sequences</i> : Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, and search.	Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci. USA</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183:63-98; and Smith, T.F. and M.S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489.	<i>ESTs</i> : fasta E value=1.06E-6 <i>Assembled ESTs</i> : fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less <i>Full Length sequences</i> : fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) <i>Nucleic Acids Res.</i> 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:88-105; and Atwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37:417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) <i>J. Mol. Biol.</i> 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) <i>Nucleic Acids Res.</i> 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score \geq GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
 - 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, 10 SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37,
 - b) a naturally occurring amino acid sequence having at least 70% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, 15 SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37,
 - 20 c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, and
 - 25 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37.
- 35 2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1,

SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:28, SEQ ID 5 NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37.

3. An isolated polynucleotide encoding a polypeptide of claim 1.
- 10 4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID 15 NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74.
- 20 6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
7. A cell transformed with a recombinant polynucleotide of claim 6.
- 25 8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method for producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide 30 comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
 - b) recovering the polypeptide so expressed.
10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

- a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74,
- 10 b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74,
- 15 c) a polynucleotide sequence complementary to a),
- d) a polynucleotide sequence complementary to b), and
- 20 e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

25 13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- 30 b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- 5 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

10

17. A composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37.

15

18. A method for treating a disease or condition associated with decreased expression of functional MEMAP, comprising administering to a patient in need of such treatment the composition of claim 16.

20

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

25

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.

20

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

30

21. A method for treating a disease or condition associated with decreased expression of functional MEMAP, comprising administering to a patient in need of such treatment a composition of claim 20.

35

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide

of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

5 23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional MEMAP, comprising administering to a patient in need of such treatment a composition of claim 23.

10 25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target 30 polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, and
- b) detecting altered expression of the target polynucleotide.

35 28. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound;
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;
- c) quantifying the amount of hybridization complex; and
- d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

SEQUENCE LISTING

<110> INCYTE GENOMICS, INC.

LAL, Preeti
YUE, Henry
TANG, Y. Tom
BANDMAN, Olga
BURFORD, Neil
AZIMZAI, Yalda
BAUGHN, Mariah R.
LU, Dyung Aina M.
PATTERSON, Chandra

<120> MEMBRANE ASSOCIATED PROTEINS

<130> PF-0731 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/149,641; 60/164,203
<151> 1999-08-17; 1999-11-09

<160> 74

<170> PERL Program

<210> 1

<211> 351

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 112301CD1

<400> 1

Met	Thr	Leu	Arg	Leu	Leu	Glu	Asp	Trp	Cys	Arg	Gly	Met	Asp	Met
1									10					15
Asn	Pro	Arg	Lys	Ala	Leu	Leu	Ile	Ala	Gly	Ile	Ser	Gln	Ser	Cys
									25					30
Ser	Val	Ala	Glu	Ile	Glu	Glu	Ala	Leu	Gln	Ala	Gly	Leu	Ala	Pro
									35		40			45
Leu	Gly	Glu	Tyr	Arg	Leu	Leu	Gly	Arg	Met	Phe	Arg	Arg	Asp	Glu
									50		55			60
Asn	Arg	Lys	Val	Ala	Leu	Val	Gly	Leu	Thr	Ala	Glu	Thr	Ser	His
									65		70			75
Ala	Leu	Val	Pro	Lys	Glu	Ile	Pro	Gly	Lys	Gly	Gly	Ile	Trp	Arg
									80		85			90
Val	Ile	Phe	Lys	Pro	Pro	Asp	Pro	Asp	Asn	Thr	Phe	Leu	Ser	Arg
									95		100			105
Leu	Asn	Glu	Phe	Leu	Ala	Gly	Glu	Gly	Met	Thr	Val	Gly	Glu	Leu
									110		115			120
Ser	Arg	Ala	Leu	Gly	His	Glu	Asn	Gly	Ser	Leu	Asp	Pro	Glu	Gln
									125		130			135
Gly	Met	Ile	Pro	Glu	Met	Trp	Ala	Pro	Met	Leu	Ala	Gln	Ala	Leu
									140		145			150
Glu	Ala	Leu	Gln	Pro	Ala	Leu	Gln	Cys	Leu	Lys	Tyr	Lys	Lys	Leu
									155		160			165
Arg	Val	Phe	Ser	Gly	Arg	Glu	Ser	Pro	Glu	Pro	Gly	Glu	Glu	Glu
									170		175			180
Phe	Gly	Arg	Trp	Met	Phe	His	Thr	Thr	Gln	Met	Ile	Lys	Ala	Trp

	185	190	195											
Gln	Val	Pro	Asp	Val	Glu	Lys	Arg	Arg	Arg	Leu	Leu	Glu	Ser	Leu
	200				205					210				
Arg	Gly	Pro	Ala	Leu	Asp	Val	Ile	Arg	Val	Leu	Lys	Ile	Asn	Asn
	215					220				225				
Pro	Leu	Ile	Thr	Val	Asp	Glu	Cys	Leu	Gln	Ala	Leu	Glu	Glu	Val
	230					235				240				
Phe	Gly	Val	Thr	Asp	Asn	Pro	Arg	Glu	Leu	Gln	Val	Lys	Tyr	Leu
	245					250				255				
Thr	Thr	Tyr	Gln	Lys	Asp	Glu	Glu	Lys	Leu	Ser	Ala	Tyr	Val	Leu
	260					265				270				
Arg	Leu	Glu	Pro	Leu	Leu	Gln	Lys	Leu	Val	Gln	Arg	Gly	Ala	Ile
	275					280				285				
Glu	Arg	Asp	Ala	Val	Asn	Gln	Ala	Arg	Leu	Asp	Gln	Val	Ile	Ala
	290					295				300				
Gly	Ala	Val	His	Lys	Thr	Ile	Arg	Arg	Glu	Leu	Asn	Leu	Pro	Glu
	305					310				315				
Asp	Gly	Pro	Ala	Pro	Gly	Phe	Leu	Gln	Leu	Leu	Val	Leu	Ile	Lys
	320					325				330				
Asp	Tyr	Glu	Ala	Ala	Glu	Glu	Glu	Ala	Leu	Leu	Gln	Ala	Ile	
	335					340				345				
Leu	Glu	Gly	Asn	Phe	Thr									
	350													

<210> 2

<211> 458

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 997947CD1

<400> 2

Met	Gln	Ala	Thr	Ser	Asn	Leu	Leu	Asn	Leu	Leu	Leu	Ser	Leu
1		5				10				15			
Phe	Ala	Gly	Leu	Asp	Pro	Ser	Lys	Thr	Gln	Ile	Ser	Pro	Lys
							20		25				30
Gly	Trp	Gln	Val	Tyr	Ser	Ser	Ala	Gln	Asp	Pro	Asp	Gly	Arg
							35		40				45
Ile	Cys	Thr	Val	Val	Ala	Pro	Glu	Gln	Asn	Leu	Cys	Ser	Arg
							50		55				60
Ala	Lys	Ser	Arg	Gln	Leu	Arg	Gln	Leu	Leu	Glu	Lys	Val	Gln
							65		70				75
Met	Ser	Gln	Ser	Ile	Glu	Val	Leu	Asn	Leu	Arg	Thr	Gln	Arg
							80		85				90
Phe	Gln	Tyr	Val	Leu	Lys	Met	Glu	Thr	Gln	Met	Lys	Gly	Leu
							95		100				105
Ala	Lys	Phe	Arg	Gln	Ile	Glu	Asp	Asp	Arg	Lys	Thr	Leu	Met
							110		115				120
Lys	His	Phe	Gln	Glu	Leu	Lys	Glu	Lys	Met	Asp	Glu	Leu	Pro
							125		130				135
Leu	Ile	Pro	Val	Leu	Glu	Gln	Tyr	Lys	Thr	Asp	Ala	Lys	Ile
							140		145				150
Thr	Gln	Phe	Lys	Glu	Glu	Ile	Arg	Asn	Leu	Ser	Ala	Val	Leu
							155		160				165
Gly	Ile	Gln	Glu	Glu	Ile	Gly	Ala	Tyr	Asp	Tyr	Glu	Glu	Leu
							170		175				180
Gln	Arg	Val	Leu	Ser	Leu	Glu	Thr	Arg	Leu	Arg	Asp	Cys	Met
							185		190				195
Lys	Leu	Thr	Cys	Gly	Lys	Leu	Met	Lys	Ile	Thr	Gly	Pro	Val
							200		205				210
Val	Lys	Thr	Ser	Gly	Thr	Arg	Phe	Gly	Ala	Trp	Met	Thr	Asp
							215		220				225

Leu Ala Ser Glu Lys Asn Asn Arg Val Trp Tyr Met Asp Ser Tyr
 230 235 240
 Thr Asn Asn Lys Ile Val Arg Glu Tyr Lys Ser Ile Ala Asp Phe
 245 250 255
 Val Ser Gly Ala Glu Ser Arg Thr Tyr Asn Leu Pro Phe Lys Trp
 260 265 270
 Ala Gly Thr Asn His Val Val Tyr Asn Gly Ser Leu Tyr Phe Asn
 275 280 285
 Lys Tyr Gln Ser Asn Ile Ile Ile Lys Tyr Ser Phe Asp Met Gly
 290 295 300
 Arg Val Leu Ala Gln Arg Ser Leu Glu Tyr Ala Gly Phe His Asn
 305 310 315
 Val Tyr Pro Tyr Thr Trp Gly Gly Phe Ser Asp Ile Asp Leu Met
 320 325 330
 Ala Asp Glu Ile Gly Leu Trp Ala Val Tyr Ala Thr Asn Gln Asn
 335 340 345
 Ala Gly Asn Ile Val Ile Ser Gln Leu Asn Gln Asp Thr Leu Glu
 350 355 360
 Val Met Lys Ser Trp Ser Thr Gly Tyr Pro Lys Arg Ser Ala Gly
 365 370 375
 Glu Ser Phe Met Ile Cys Gly Thr Leu Tyr Val Thr Asn Ser His
 380 385 390
 Leu Thr Gly Ala Lys Val Tyr Tyr Ser Tyr Ser Thr Lys Thr Ser
 395 400 405
 Thr Tyr Glu Tyr Thr Asp Ile Pro Phe His Asn Gln Tyr Phe His
 410 415 420
 Ile Ser Met Leu Asp Tyr Asn Ala Arg Asp Arg Ala Leu Tyr Ala
 425 430 435
 Trp Asn Asn Gly His Gln Val Leu Phe Asn Val Thr Leu Phe His
 440 445 450
 Ile Ile Lys Thr Glu Asp Asp Thr
 455

<210> 3

<211> 219

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1521513CD1

<400> 3

Met Asn Ser Ser Lys Ser Ser Glu Thr Gln Cys Thr Glu Arg Gly
 1 5 10 15
 Cys Phe Ser Ser Gln Met Phe Leu Trp Thr Val Ala Gly Ile Pro
 20 25 30
 Ile Leu Phe Leu Ser Ala Cys Phe Ile Thr Arg Cys Val Val Thr
 35 40 45
 Phe Arg Ile Phe Gln Thr Cys Asp Glu Lys Lys Phe Gln Leu Pro
 50 55 60
 Glu Asn Phe Thr Glu Leu Ser Cys Tyr Asn Tyr Gly Ser Gly Ser
 65 70 75
 Val Lys Asn Cys Cys Pro Leu Asn Trp Glu Tyr Phe Gln Ser Ser
 80 85 90
 Cys Tyr Phe Phe Ser Thr Asp Thr Ile Ser Trp Ala Leu Ser Leu
 95 100 105
 Lys Asn Cys Ser Ala Met Gly Ala His Leu Val Val Ile Asn Ser
 110 115 120
 Gln Glu Glu Gln Glu Phe Leu Ser Tyr Lys Lys Pro Lys Met Arg
 125 130 135
 Glu Phe Phe Ile Gly Leu Ser Asp Gln Val Val Glu Gly Gln Trp
 140 145 150
 Gln Trp Val Asp Gly Thr Pro Leu Thr Lys Ser Leu Ser Phe Trp

Asp Val Gly Glu	155	Pro Asn Asn Ile Ala	160	Thr Leu Glu Asp Cys	165
170		175		Ala	180
Thr Met Arg Asp		Ser Ser Asn Pro Arg		Gln Asn Trp Asn Asp	
185		190		Val	195
Thr Cys Phe Leu Asn		Tyr Phe Arg Ile		Cys Glu Met Val Gly	Ile
200		205		210	
Asn Pro Leu Asn Lys		Gly Lys Ser Leu			
215					

<210> 4

<211> 276

<212> PRT

<213> Homo sapiens

<220>

!

<221> misc_feature

<223> Incyte ID No: 1863994CD1

<400> 4

Met Glu Ser Arg Met Trp Pro Ala Leu Leu Leu Ser His Leu Leu					
1	5	10	15		
Pro Leu Trp Pro Leu Leu Leu Pro Leu Pro Pro Pro Ala Gln					
20	25	30			
Gly Ser Ser Ser Pro Arg Thr Pro Pro Ala Pro Ala Arg Pro					
35	40	45			
Pro Cys Ala Arg Gly Gly Pro Ser Ala Pro Arg His Val Cys Val					
50	55	60			
Trp Glu Arg Ala Pro Pro Pro Ser Arg Ser Pro Arg Val Pro Arg					
65	70	75			
Ser Arg Arg Gln Val Leu Pro Gly Thr Ala Pro Pro Ala Thr Pro					
80	85	90			
Ser Gly Phe Glu Glu Gly Pro Pro Ser Ser Gln Tyr Pro Trp Ala					
95	100	105			
Ile Val Trp Gly Pro Thr Val Ser Arg Glu Asp Gly Gly Asp Pro					
110	115	120			
Asn Ser Ala Asn Pro Gly Phe Leu Asp Tyr Gly Phe Ala Ala Pro					
125	130	135			
His Gly Leu Ala Thr Pro His Pro Asn Ser Asp Ser Met Arg Gly					
140	145	150			
Asp Gly Asp Gly Leu Ile Leu Gly Glu Ala Pro Ala Thr Leu Arg					
155	160	165			
Pro Phe Leu Phe Gly Gly Arg Gly Glu Gly Val Asp Pro Gln Leu					
170	175	180			
Tyr Val Thr Ile Thr Ile Ser Ile Ile Ile Val Leu Val Ala Thr					
185	190	195			
Gly Ile Ile Phe Lys Phe Cys Trp Asp Arg Ser Gln Lys Arg Arg					
200	205	210			
Arg Pro Ser Gly Gln Gln Gly Ala Leu Arg Gln Glu Glu Ser Gln					
215	220	225			
Gln Pro Leu Thr Asp Leu Ser Pro Ala Gly Val Thr Val Leu Gly					
230	235	240			
Ala Phe Gly Asp Ser Pro Thr Pro Thr Pro Asp His Glu Glu Pro					
245	250	255			
Arg Gly Gly Pro Arg Pro Gly Met Pro His Pro Lys Gly Ala Pro					
260	265	270			
Ala Phe Gln Leu Asn Arg					
275					

<210> 5

<211> 375

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature
 <223> Incyte ID No: 2071941CD1

<400> 5
 Met Ser Ser His Lys Gly Ser Val Val Ala Gln Gly Asn Gly Ala
 1 5 10 15
 Pro Ala Ser Asn Arg Glu Ala Asp Thr Val Glu Leu Ala Glu Leu
 20 25 30
 Gly Pro Leu Leu Glu Glu Lys Gly Lys Arg Val Ile Ala Asn Pro
 35 40 45
 Pro Lys Ala Glu Glu Gln Thr Cys Pro Val Pro Gln Glu Glu
 50 55 60
 Glu Glu Glu Val Arg Val Leu Thr Leu Pro Leu Gln Ala His His
 65 70 75
 Ala Met Glu Lys Met Glu Glu Phe Val Tyr Lys Val Trp Glu Gly
 80 85 90
 Arg Trp Arg Val Ile Pro Tyr Asp Val Leu Pro Asp Trp Leu Lys
 95 100 105
 Asp Asn Asp Tyr Leu Leu His Gly His Arg Pro Pro Met Pro Ser
 110 115 120
 Phe Arg Ala Cys Phe Lys Ser Ile Phe Arg Ile His Thr Glu Thr
 125 130 135
 Gly Asn Ile Trp Thr His Leu Leu Gly Phe Val Leu Phe Leu Phe
 140 145 150
 Leu Gly Ile Leu Thr Met Leu Arg Pro Asn Met Tyr Phe Met Ala
 155 160 165
 Pro Leu Gln Glu Lys Val Val Phe Gly Met Phe Leu Gly Ala
 170 175 180
 Val Leu Cys Leu Ser Phe Ser Trp Leu Phe His Thr Val Tyr Cys
 185 190 195
 His Ser Glu Lys Val Ser Arg Thr Phe Ser Lys Leu Asp Tyr Ser
 200 205 210
 Gly Ile Ala Leu Leu Ile Met Gly Ser Phe Val Pro Trp Leu Tyr
 215 220 225
 Tyr Ser Phe Tyr Cys Ser Pro Gln Pro Arg Leu Ile Tyr Leu Ser
 230 235 240
 Ile Val Cys Val Leu Gly Ile Ser Ala Ile Ile Val Ala Gln Trp
 245 250 255
 Asp Arg Phe Ala Thr Pro Lys His Arg Gln Thr Arg Ala Gly Val
 260 265 270
 Phe Leu Gly Leu Gly Leu Ser Gly Val Val Pro Thr Met His Phe
 275 280 285
 Thr Ile Ala Glu Gly Phe Val Lys Ala Thr Thr Val Gly Gln Met
 290 295 300
 Gly Trp Phe Phe Leu Met Ala Val Met Tyr Ile Thr Gly Ala Gly
 305 310 315
 Leu Tyr Ala Ala Arg Ile Pro Glu Arg Phe Phe Pro Gly Lys Phe
 320 325 330
 Asp Ile Trp Phe Gln Ser His Gln Ile Phe His Val Leu Val Val
 335 340 345
 Ala Ala Ala Phe Val His Phe Tyr Gly Val Ser Asn Leu Gln Glu
 350 355 360
 Phe Arg Tyr Gly Leu Glu Gly Gly Cys Thr Asp Asp Thr Leu Leu
 365 370 375

<210> 6
 <211> 249
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2172512CD1

<400> 6

Met	Ser	Gly	Val	Val	Pro	Thr	Ala	Pro	Glu	Gln	Pro	Ala	Gly	Glu
1			5					10						15
Met	Glu	Asn	Gln	Thr	Lys	Pro	Pro	Asp	Pro	Arg	Pro	Asp	Ala	Pro
				20				25						30
Pro	Glu	Tyr	Ser	Ser	His	Phe	Leu	Pro	Gly	Pro	Pro	Gly	Thr	Ala
					35			40						45
Val	Pro	Pro	Pro	Thr	Gly	Tyr	Pro	Gly	Gly	Leu	Pro	Met	Gly	Tyr
				50				55						60
Tyr	Ser	Pro	Gln	Gln	Pro	Ser	Thr	Phe	Pro	Leu	Tyr	Gln	Pro	Val
				65				70						75
Gly	Gly	Ile	His	Pro	Val	Arg	Tyr	Gln	Pro	Gly	Lys	Tyr	Pro	Met
				80				85						90
Pro	Asn	Gln	Ser	Val	Pro	Ile	Thr	Trp	Met	Pro	Gly	Pro	Thr	Pro
				95				100						105
Met	Ala	Asn	Cys	Pro	Pro	Gly	Leu	Glu	Tyr	Leu	Val	Gln	Leu	Asp
				110				115						120
Asn	Ile	His	Val	Leu	Gln	His	Phe	Glu	Pro	Leu	Glu	Met	Met	Thr
				125				130						135
Cys	Phe	Glu	Thr	Asn	Asn	Arg	Tyr	Asp	Ile	Lys	Asn	Asn	Ser	Asp
				140				145						150
Gln	Met	Val	Tyr	Ile	Val	Thr	Glu	Asp	Thr	Asp	Asp	Phe	Thr	Arg
				155				160						165
Asn	Ala	Tyr	Arg	Thr	Leu	Arg	Pro	Phe	Val	Leu	Arg	Val	Thr	Asp
				170				175						180
Cys	Met	Gly	Arg	Glu	Ile	Met	Thr	Met	Gln	Arg	Pro	Phe	Arg	Cys
				185				190						195
Thr	Cys	Cys	Cys	Phe	Cys	Cys	Pro	Ser	Ala	Arg	Gln	Glu	Leu	Glu
				200				205						210
Val	Gln	Cys	Pro	Pro	Gly	Val	Thr	Ile	Gly	Phe	Val	Ala	Glu	His
				215				220						225
Trp	Asn	Leu	Cys	Arg	Ala	Val	Tyr	Ser	Ile	Gln	Lys	Lys	Lys	Lys
				230				235						240
Lys	Ile	Ala	Ala	Gln	Ala	Tyr	Ser	Leu						
				245										

<210> 7

<211> 353

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2483172CD1

<400> 7

Met	Ala	Met	Thr	Leu	Leu	Glu	Asp	Trp	Cys	Arg	Gly	Met	Asp	Val
1		5				10								15
Asn	Ser	Gln	Arg	Ala	Leu	Leu	Val	Trp	Gly	Ile	Pro	Val	Asn	Cys
				20				25						30
Asp	Glu	Ala	Glu	Ile	Glu	Glu	Thr	Leu	Gln	Ala	Ala	Met	Pro	Gln
				35				40						45
Val	Ser	Tyr	Arg	Met	Leu	Gly	Arg	Met	Phe	Trp	Arg	Glu	Glu	Asn
				50				55						60
Ala	Lys	Ala	Ala	Leu	Leu	Glu	Leu	Thr	Gly	Ala	Val	Asp	Tyr	Ala
				65				70						75
Ala	Ile	Pro	Arg	Glu	Met	Pro	Gly	Lys	Gly	Gly	Val	Trp	Lys	Val
				80				85						90
Leu	Phe	Lys	Pro	Pro	Thr	Ser	Asp	Ala	Glu	Phe	Leu	Glu	Arg	Leu
				95				100						105
His	Leu	Phe	Leu	Ala	Arg	Glu	Gly	Trp	Thr	Val	Gln	Asp	Val	Ala
				110				115						120
Arg	Val	Leu	Gly	Phe	Gln	Asn	Pro	Thr	Pro	Thr	Pro	Gly	Pro	Glu
				125				130						135

Met Pro Ala Glu Met Leu Asn Tyr Ile Leu Asp Asn Val Ile Gln
 140 145 150
 Pro Leu Val Glu Ser Ile Trp Tyr Lys Arg Leu Thr Leu Phe Ser
 155 160 165
 Gly Arg Asp Ile Pro Gly Pro Gly Glu Glu Thr Phe Asp Pro Trp
 170 175 180
 Leu Glu His Thr Asn Glu Val Leu Glu Glu Trp Gln Val Ser Asp
 185 190 195
 Val Glu Lys Arg Arg Arg Leu Met Glu Ser Leu Arg Gly Pro Ala
 200 205 210
 Ala Asp Val Ile Arg Ile Leu Lys Ser Asn Asn Pro Ala Ile Thr
 215 220 225
 Thr Ala Glu Cys Leu Lys Ala Leu Glu Gln Val Phe Gly Ser Val
 230 235 240
 Glu Ser Ser Arg Asp Ala Gln Ile Lys Phe Leu Asn Thr Tyr Gln
 245 250 255
 Asn Pro Gly Glu Lys Leu Ser Ala Tyr Val Ile Arg Leu Glu Pro
 260 265 270
 Leu Leu Gln Lys Val Val Glu Lys Gly Ala Ile Asp Lys Asp Asn
 275 280 285
 Val Asn Gln Ala Arg Leu Glu Gln Val Ile Ala Gly Ala Asn His
 290 295 300
 Ser Gly Ala Ile Arg Arg Gln Leu Trp Leu Thr Gly Ala Gly Glu
 305 310 315
 Gly Pro Ala Pro Asn Leu Phe Gln Leu Leu Val Gln Ile Arg Glu
 320 325 330
 Glu Glu Ala Lys Glu Glu Glu Glu Ala Glu Ala Thr Leu Leu
 335 340 345
 Gln Leu Gly Leu Glu Gly His Phe
 350

<210> 8

<211> 194

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2656128CD1

<400> 8

Met His Asp Ser Asn Asn Val Glu Lys Asp Ile Thr Pro Ser Glu
 1 5 10 15
 Leu Pro Ala Asn Pro Gly Cys Leu His Ser Lys Glu His Ser Ile
 20 25 30
 Lys Ala Thr Leu Ile Trp Arg Leu Phe Phe Leu Ile Met Phe Leu
 35 40 45
 Thr Ile Ile Val Cys Gly Met Val Ala Ala Leu Ser Ala Ile Arg
 50 55 60
 Ala Asn Cys His Gln Glu Pro Ser Val Cys Leu Gln Ala Ala Cys
 65 70 75
 Pro Glu Ser Trp Ile Gly Phe Gln Arg Lys Cys Phe Tyr Phe Ser
 80 85 90
 Asp Asp Thr Lys Asn Trp Thr Ser Ser Gln Arg Phe Cys Asp Ser
 95 100 105
 Gln Asp Ala Asp Leu Ala Gln Val Glu Ser Phe Gln Glu Leu Asn
 110 115 120
 Phe Leu Leu Arg Tyr Lys Gly Pro Ser Asp His Trp Ile Gly Leu
 125 130 135
 Ser Arg Glu Gln Gly Gln Pro Trp Lys Trp Ile Asn Gly Thr Glu
 140 145 150
 Trp Thr Arg Gln Leu Val Met Lys Glu Asp Gly Ala Asn Leu Tyr
 155 160 165
 Val Ala Lys Val Ser Gln Val Pro Arg Met Asn Pro Arg Pro Val

170	175	180
Met Val Ser Tyr Pro Gly Ser Arg Arg	Val Cys Leu Phe Glu	
185	190	
<210> 9		
<211> 322		
<212> PRT		
<213> Homo sapiens		
<220>		
<221> misc_feature		
<223> Incyte ID No: 5855841CD1		
<400> 9		
Met Ser Ser Leu Gly Gly Ser Gln Asp Ala Gly Gly Ser Ser		
1 5 10 15		
Ser Ser Ser Thr Asn Gly Ser Gly Gly Ser Gly Ser Ser Gly Pro		
20 25 30		
Lys Ala Gly Ala Ala Asp Lys Ser Ala Val Val Ala Ala Ala		
35 40 45		
Pro Ala Ser Val Ala Asp Asp Thr Pro Pro Pro Glu Arg Arg Asn		
50 55 60		
Lys Ser Gly Ile Ile Ser Glu Pro Leu Asn Lys Ser Leu Arg Arg		
65 70 75		
Ser Arg Pro Leu Ser His Tyr Ser Ser Phe Gly Ser Ser Gly Gly		
80 85 90		
Ser Gly Gly Ser Met Met Gly Gly Glu Ser Ala Asp Lys Ala		
95 100 105		
Thr Ala Ala Ala Ala Ala Ser Leu Leu Ala Asn Gly His Asp		
110 115 120		
Leu Ala Ala Ala Met Ala Val Asp Lys Ser Asn Pro Thr Ser Lys		
125 130 135		
His Lys Ser Gly Ala Val Ala Ser Leu Leu Ser Lys Ala Glu Arg		
140 145 150		
Ala Thr Glu Leu Ala Ala Glu Gly Gln Leu Thr Leu Gln Gln Phe		
155 160 165		
Ala Gln Ser Thr Glu Met Leu Lys Arg Val Val Gln Glu His Leu		
170 175 180		
Pro Leu Met Ser Glu Ala Gly Ala Gly Leu Pro Asp Met Glu Ala		
185 190 195		
Val Ala Gly Ala Glu Ala Leu Asn Gly Gln Ser Asp Phe Pro Tyr		
200 205 210		
Leu Gly Ala Phe Pro Ile Asn Pro Gly Leu Phe Ile Met Thr Pro		
215 220 225		
Ala Gly Val Phe Leu Ala Glu Ser Ala Leu His Met Ala Gly Leu		
230 235 240		
Ala Glu Tyr Pro Met Gln Gly Glu Leu Ala Ser Ala Ile Ser Ser		
245 250 255		
Gly Lys Lys Lys Arg Lys Arg Cys Gly Met Cys Ala Pro Cys Arg		
260 265 270		
Arg Arg Ile Asn Cys Glu Gln Cys Ser Ser Cys Arg Asn Arg Lys		
275 280 285		
Thr Gly His Gln Ile Cys Lys Phe Arg Lys Cys Glu Glu Leu Lys		
290 295 300		
Lys Lys Pro Ser Ala Ala Leu Glu Lys Val Met Leu Pro Thr Gly		
305 310 315		
Ala Ala Phe Arg Trp Phe Gln		
320		

<210> 10

<211> 335

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature
 <223> Incyte ID No: 603462CD1

<400> 10
 Met Leu Gln Gly His Ser Ser Val Phe Gln Ala Leu Leu Gly Thr
 1 5 10 15
 Phe Phe Thr Trp Gly Met Thr Ala Ala Gly Ala Ala Leu Val Phe
 20 25 30
 Val Phe Ser Ser Gly Gln Arg Arg Ile Leu Asp Gly Ser Leu Gly
 35 40 45
 Phe Ala Ala Gly Val Met Leu Ala Ala Ser Tyr Trp Ser Leu Leu
 50 55 60
 Ala Pro Ala Val Glu Met Ala Thr Ser Ser Gly Gly Phe Gly Ala
 65 70 75
 Phe Ala Phe Phe Pro Val Ala Val Gly Phe Thr Leu Gly Ala Ala
 80 85 90
 Phe Val Tyr Leu Ala Asp Leu Leu Met Pro His Leu Gly Ala Ala
 95 100 105
 Glu Asp Pro Gln Thr Ala Leu Ala Leu Asn Phe Gly Ser Thr Leu
 110 115 120
 Met Lys Lys Lys Ser Asp Pro Glu Gly Pro Ala Leu Leu Phe Pro
 125 130 135
 Glu Ser Glu Leu Ser Ile Arg Ile Asp Lys Ser Glu Asn Gly Glu
 140 145 150
 Ala Tyr Gln Arg Lys Lys Ala Ala Ala Thr Gly Leu Pro Glu Gly
 155 160 165
 Pro Ala Val Pro Val Pro Ser Arg Gly Asn Leu Ala Gln Pro Gly
 170 175 180
 Gly Ser Ser Trp Arg Arg Ile Ala Leu Ile Leu Ala Ile Thr
 185 190 195
 Ile His Asn Val Pro Glu Gly Leu Ala Val Gly Val Gly Phe Gly
 200 205 210
 Ala Ile Glu Lys Thr Ala Ser Ala Thr Phe Glu Ser Ala Arg Asn
 215 220 225
 Leu Ala Ile Gly Ile Gly Ile Gln Asn Phe Pro Glu Gly Leu Ala
 230 235 240
 Val Ser Leu Pro Leu Arg Gly Ala Gly Phe Ser Thr Trp Arg Ala
 245 250 255
 Phe Trp Tyr Gly Gln Leu Ser Gly Met Val Glu Pro Leu Ala Gly
 260 265 270
 Val Phe Gly Ala Phe Ala Val Val Leu Ala Glu Pro Ile Leu Pro
 275 280 285
 Tyr Ala Leu Ala Phe Ala Ala Gly Ala Met Val Tyr Val Val Met
 290 295 300
 Asp Asp Ile Ile Pro Glu Ala Gln Ile Ser Gly Asn Gly Lys Leu
 305 310 315
 Ala Ser Trp Ala Ser Ile Leu Gly Phe Val Val Met Met Ser Leu
 320 325 330
 Asp Val Gly Leu Gly
 335

<210> 11

<211> 620

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 747681CD1

<400> 11

Met Gln Val Ser Lys Arg Met Leu Ala Gly Gly Val Arg Ser Met
 1 5 10 15
 Pro Ser Pro Leu Leu Ala Cys Trp Gln Pro Ile Leu Leu Val

20	25	30												
Leu	Gly	Ser	Val	Leu	Ser	Gly	Ser	Ala	Thr	Gly	Cys	Pro	Pro	Arg
35						40								45
Cys	Glu	Cys	Ser	Ala	Gln	Asp	Arg	Ala	Val	Leu	Cys	His	Arg	Lys
50							55							60
Arg	Phe	Val	Ala	Val	Pro	Glu	Gly	Ile	Pro	Thr	Glu	Thr	Arg	Leu
65								70						75
Leu	Asp	Leu	Gly	Lys	Asn	Arg	Ile	Lys	Thr	Leu	Asn	Gln	Asp	Glu
80								85						90
Phe	Ala	Ser	Phe	Pro	His	Leu	Glu	Glu	Leu	Glu	Leu	Asn	Glu	Asn
95								100						105
Ile	Val	Ser	Ala	Val	Glu	Pro	Gly	Ala	Phe	Asn	Asn	Leu	Phe	Asn
110								115						120
Leu	Arg	Thr	Leu	Gly	Leu	Arg	Ser	Asn	Arg	Leu	Lys	Leu	Ile	Pro
125								130						135
Leu	Gly	Val	Phe	Thr	Gly	Leu	Ser	Asn	Leu	Thr	Lys	Leu	Asp	Ile
140								145						150
Ser	Glu	Asn	Lys	Ile	Val	Ile	Leu	Leu	Asp	Tyr	Met	Phe	Gln	Asp
155								160						165
Leu	Tyr	Asn	Leu	Lys	Ser	Leu	Glu	Val	Gly	Asp	Asn	Asp	Leu	Val
170								175						180
Tyr	Ile	Ser	His	Arg	Ala	Phe	Ser	Gly	Leu	Asn	Ser	Leu	Glu	Gln
185								190						195
Leu	Thr	Leu	Glu	Lys	Cys	Asn	Leu	Thr	Ser	Ile	Pro	Thr	Glu	Ala
200								205						210
Leu	Ser	His	Leu	His	Gly	Leu	Ile	Val	Leu	Arg	Leu	Arg	His	Leu
215								220						225
Asn	Ile	Asn	Ala	Ile	Arg	Asp	Tyr	Ser	Phe	Lys	Arg	Leu	Tyr	Arg
230								235						240
Leu	Lys	Val	Leu	Glu	Ile	Ser	His	Trp	Pro	Tyr	Leu	Asp	Thr	Met
245								250						255
Thr	Pro	Asn	Cys	Leu	Tyr	Gly	Leu	Asn	Leu	Thr	Ser	Leu	Ser	Ile
260								265						270
Thr	His	Cys	Asn	Leu	Thr	Ala	Val	Pro	Tyr	Leu	Ala	Val	Arg	His
275								280						285
Leu	Val	Tyr	Leu	Arg	Phe	Leu	Asn	Leu	Ser	Tyr	Asn	Pro	Ile	Ser
290								295						300
Thr	Ile	Glu	Gly	Ser	Met	Leu	His	Glu	Leu	Leu	Arg	Leu	Gln	Glu
305								310						315
Ile	Gln	Leu	Val	Gly	Gly	Gln	Leu	Ala	Val	Val	Glu	Pro	Tyr	Ala
320								325						330
Phe	Arg	Gly	Leu	Asn	Tyr	Leu	Arg	Val	Leu	Asn	Val	Ser	Gly	Asn
335								340						345
Gln	Leu	Thr	Thr	Leu	Glu	Glu	Ser	Val	Phe	His	Ser	Val	Gly	Asn
350								355						360
Leu	Glu	Thr	Leu	Ile	Leu	Asp	Ser	Asn	Pro	Leu	Ala	Cys	Asp	Cys
365								370						375
Arg	Leu	Leu	Trp	Val	Phe	Arg	Arg	Arg	Trp	Arg	Leu	Asn	Phe	Asn
380								385						390
Arg	Gln	Gln	Pro	Thr	Cys	Ala	Thr	Pro	Glu	Phe	Val	Gln	Gly	Lys
395								400						405
Glu	Phe	Lys	Asp	Phe	Pro	Asp	Val	Leu	Leu	Pro	Asn	Tyr	Phe	Thr
410								415						420
Cys	Arg	Arg	Ala	Arg	Ile	Arg	Asp	Arg	Lys	Ala	Gln	Gln	Val	Phe
425								430						435
Val	Asp	Glu	Gly	His	Thr	Val	Gln	Phe	Val	Cys	Arg	Ala	Asp	Gly
440								445						450
Asp	Pro	Pro	Pro	Ala	Ile	Leu	Trp	Leu	Ser	Pro	Arg	Lys	His	Leu
455								460						465
Val	Ser	Ala	Lys	Ser	Asn	Gly	Arg	Leu	Thr	Val	Phe	Pro	Asp	Gly
470								475						480
Thr	Leu	Glu	Val	Arg	Tyr	Ala	Gln	Val	Gln	Asp	Asn	Gly	Thr	Tyr
485								490						495

Leu Cys Ile Ala Ala Asn Ala Gly Gly Asn Asp Ser Met Pro Ala
 500 505 510
 His Leu His Val Arg Ser Tyr Ser Pro Asp Trp Pro His Gln Pro
 515 520 525
 Asn Lys Thr Phe Ala Phe Ile Ser Asn Gln Pro Gly Glu Gly Glu
 530 535 540
 Ala Asn Ser Thr Arg Ala Thr Val Pro Phe Pro Phe Asp Ile Lys
 545 550 555
 Thr Leu Ile Ile Ala Thr Thr Met Gly Phe Ile Ser Phe Leu Gly
 560 565 570
 Val Val Leu Phe Cys Leu Val Leu Leu Phe Leu Trp Ser Arg Gly
 575 580 585
 Lys Gly Asn Thr Lys His Asn Ile Glu Ile Glu Tyr Val Pro Arg
 590 595 600
 Lys Ser Asp Ala Gly Ile Ser Ser Ala Asp Ala Pro Arg Lys Phe
 605 610 615
 Asn Met Lys Met Ile
 620

<210> 12
 <211> 491
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 919469CD1

<400> 12

Met Ala Gly Gln Gly Leu Pro Leu His Val Ala Thr Leu Leu Thr	1	5	10	15
Gly Leu Leu Glu Cys Leu Gly Phe Ala Gly Val Leu Phe Gly Trp	20	25	30	
Pro Ser Leu Val Phe Val Phe Lys Asn Glu Asp Tyr Phe Lys Asp	35	40	45	
Leu Cys Gly Pro Asp Ala Gly Pro Ile Gly Asn Ala Thr Gly Gln	50	55	60	
Ala Asp Cys Lys Ala Gln Asp Glu Arg Phe Ser Leu Ile Phe Thr	65	70	75	
Leu Gly Ser Phe Met Asn Asn Phe Met Thr Phe Pro Thr Gly Tyr	80	85	90	
Ile Phe Asp Arg Phe Lys Thr Thr Val Ala Arg Leu Ile Ala Ile	95	100	105	
Phe Phe Tyr Thr Thr Ala Thr Leu Ile Ile Ala Phe Thr Ser Ala	110	115	120	
Gly Ser Ala Val Leu Leu Phe Leu Ala Met Pro Met Leu Thr Ile	125	130	135	
Gly Gly Ile Leu Phe Leu Ile Thr Asn Leu Gln Ile Gly Asn Leu	140	145	150	
Phe Gly Gln His Arg Ser Thr Ile Ile Thr Leu Tyr Asn Gly Ala	155	160	165	
Phe Asp Ser Ser Ser Ala Val Phe Leu Ile Ile Lys Leu Leu Tyr	170	175	180	
Glu Lys Gly Ile Ser Leu Arg Ala Ser Phe Ile Phe Ile Ser Val	185	190	195	
Cys Ser Thr Trp His Val Ala Arg Thr Phe Leu Leu Met Pro Arg	200	205	210	
Gly His Ile Pro Tyr Pro Leu Pro Pro Asn Tyr Ser Tyr Gly Leu	215	220	225	
Cys Pro Gly Asn Gly Thr Thr Lys Glu Glu Lys Glu Thr Ala Glu	230	235	240	
His Glu Asn Arg Glu Leu Gln Ser Lys Glu Phe Leu Ser Ala Lys	245	250	255	
Glu Glu Thr Pro Gly Ala Gly Gln Lys Gln Glu Leu Arg Ser Phe				

260	265	270
Trp Ser Tyr Ala Phe Ser Arg Arg Phe	Ala Trp His Leu Val	Trp
275	280	285
Leu Ser Val Ile Gln Leu Trp His Tyr	Leu Phe Ile Gly Thr	Leu
290	295	300
Asn Ser Leu Leu Thr Asn Met Ala Gly	Gly Asp Met Ala Arg	Val
305	310	315
Ser Thr Tyr Thr Asn Ala Phe Ala Phe	Thr Gln Phe Gly Val	Leu
320	325	330
Cys Ala Pro Trp Asn Gly Leu Leu Met	Asp Arg Leu Lys Gln	Lys
335	340	345
Tyr Gln Lys Glu Ala Arg Lys Thr Gly	Ser Ser Thr Leu Ala	Val
350	355	360
Ala Leu Cys Ser Thr Val Pro Ser Leu	Ala Leu Thr Ser Leu	Leu
365	370	375
Cys Leu Gly Phe Ala Leu Cys Ala Ser	Val Pro Ile Leu Pro	Leu
380	385	390
Gln Tyr Leu Thr Phe Ile Leu Gln Val	Ile Ser Arg Ser Phe	Leu
395	400	405
Tyr Gly Ser Asn Ala Ala Phe Leu Thr	Leu Ala Phe Pro Ser	Glu
410	415	420
His Phe Gly Lys Leu Phe Gly Leu Val	Met Ala Leu Ser Ala	Val
425	430	435
Val Ser Leu Leu Gln Phe Pro Ile Phe	Thr Leu Ile Lys Gly	Ser
440	445	450
Leu Gln Asn Asp Pro Phe Tyr Val Asn	Val Met Phe Met Leu	Ala
455	460	465
Ile Leu Leu Thr Phe Phe His Pro Phe	Leu Val Tyr Arg Glu	Cys
470	475	480
Arg Thr Trp Lys Glu Ser Pro Ser Ala	Ile Ala	
485	490	

<210> 13

<211> 580

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 977658CD1

<400> 13

Met Thr Ala Pro Ala Gly Pro Arg Gly	Ser Glu Thr Glu Arg	Leu	
1	5	10	15
Leu Thr Pro Asn Pro Gly Tyr Gly	Thr Gln Ala Gly Pro Ser	Pro	
20	25	30	
Ala Pro Pro Thr Pro Pro Glu Glu	Asp Leu Arg Arg Arg	Leu	
35	40	45	
Lys Tyr Phe Phe Met Ser Pro Cys Asp	Lys Phe Arg Ala Lys	Gly	
50	55	60	
Arg Lys Pro Cys Lys Leu Met Leu	Gln Val Val Lys Ile Leu	Val	
65	70	75	
Val Thr Val Gln Leu Ile Leu Phe	Gly Leu Ser Asn Gln	Leu Ala	
80	85	90	
Val Thr Phe Arg Glu Glu Asn Thr	Ile Ala Phe Arg His	Leu Phe	
95	100	105	
Leu Leu Gly Tyr Ser Asp Gly Ala Asp	Asp Thr Phe Ala Ala	Tyr	
110	115	120	
Thr Arg Glu Gln Leu Tyr Gln Ala	Ile Phe His Ala Val	Asp Gln	
125	130	135	
Tyr Leu Ala Leu Pro Asp Val Ser	Leu Gly Arg Tyr Ala	Tyr Val	
140	145	150	
Arg Gly Gly Gly Asp Pro Trp Thr Asn	Gly Ser Gly Leu Ala	Leu	
155	160	165	

Cys Gln Arg Tyr Tyr His Arg Gly His Val Asp Pro Ala Asn Asp
 170 175 180
 Thr Phe Asp Ile Asp Pro Met Val Val Thr Asp Cys Ile Gln Val
 185 190 195
 Asp Pro Pro Glu Arg Pro Pro Pro Pro Ser Asp Asp Leu Thr
 200 205 210
 Leu Leu Glu Ser Ser Ser Tyr Lys Asn Leu Thr Leu Lys Phe
 215 220 225
 His Lys Leu Val Asn Val Thr Ile His Phe Arg Leu Lys Thr Ile
 230 235 240
 Asn Leu Gln Ser Leu Ile Asn Asn Glu Ile Pro Asp Cys Tyr Thr
 245 250 255
 Phe Ser Val Leu Ile Thr Phe Asp Asn Lys Ala His Ser Gly Arg
 260 265 270
 Ile Pro Ile Ser Leu Glu Thr Gln Ala His Ile Gln Glu Cys Lys
 275 280 285
 His Pro Ser Val Phe Gln His Gly Asp Asn Ser Phe Arg Leu Leu
 290 295 300
 Phe Asp Val Val Val Ile Leu Thr Cys Ser Leu Ser Phe Leu Leu
 305 310 315
 Cys Ala Arg Ser Leu Leu Arg Gly Phe Leu Leu Gln Asn Glu Phe
 320 325 330
 Val Gly Phe Met Trp Arg Gln Arg Gly Arg Val Ile Ser Leu Trp
 335 340 345
 Glu Arg Leu Glu Phe Val Asn Gly Trp Tyr Ile Leu Leu Val Thr
 350 355 360
 Ser Asp Val Leu Thr Ile Ser Gly Thr Ile Met Lys Ile Gly Ile
 365 370 375
 Glu Ala Lys Asn Leu Ala Ser Tyr Asp Val Cys Ser Ile Leu Leu
 380 385 390
 Gly Thr Ser Thr Leu Leu Val Trp Val Gly Val Ile Arg Tyr Leu
 395 400 405
 Thr Phe Phe His Asn Tyr Asn Ile Leu Ile Ala Thr Leu Arg Val
 410 415 420
 Ala Leu Pro Ser Val Met Arg Phe Cys Cys Val Ala Val Ile
 425 430 435
 Tyr Leu Gly Tyr Cys Phe Cys Gly Trp Ile Val Leu Gly Pro Tyr
 440 445 450
 His Val Lys Phe Arg Ser Leu Ser Met Val Ser Glu Cys Leu Phe
 455 460 465
 Ser Leu Ile Asn Gly Asp Asp Met Phe Val Thr Phe Ala Ala Met
 470 475 480
 Gln Ala Gln Gln Gly Arg Ser Ser Leu Val Trp Leu Phe Ser Gln
 485 490 495
 Leu Tyr Leu Tyr Ser Phe Ile Ser Leu Phe Ile Tyr Met Val Leu
 500 505 510
 Ser Leu Phe Ile Ala Leu Ile Thr Gly Ala Tyr Asp Thr Ile Lys
 515 520 525
 His Pro Gly Gly Ala Gly Ala Glu Glu Ser Glu Leu Gln Ala Tyr
 530 535 540
 Ile Ala Gln Cys Gln Asp Ser Pro Thr Ser Gly Lys Phe Arg Arg
 545 550 555
 Gly Ser Gly Ser Ala Cys Ser Leu Leu Cys Cys Cys Gly Arg Asp
 560 565 570
 Pro Ser Glu Glu His Ser Leu Leu Val Asn
 575 580

<210> 14

<211> 455

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1004703CD1

<400> 14

Met	Ser	Phe	Leu	Ile	Asp	Ser	Ser	Ile	Met	Ile	Thr	Ser	Gln	Ile
1		5						10					15	
Leu	Phe	Phe	Gly	Phe	Gly	Trp	Leu	Phe	Phe	Met	Arg	Gln	Leu	Phe
							20			25			30	
Lys	Asp	Tyr	Glu	Ile	Arg	Gln	Tyr	Val	Val	Gln	Val	Ile	Phe	Ser
							35			40			45	
Val	Thr	Phe	Ala	Phe	Ser	Cys	Thr	Met	Phe	Glu	Leu	Ile	Ile	Phe
							50			55			60	
Glu	Ile	Leu	Gly	Val	Leu	Asn	Ser	Ser	Ser	Arg	Tyr	Phe	His	Trp
							65			70			75	
Lys	Met	Asn	Leu	Cys	Val	Ile	Leu	Ile	Leu	Val	Phe	Met	Val	
							80			85			90	
Pro	Phe	Tyr	Ile	Gly	Tyr	Phe	Ile	Val	Ser	Asn	Ile	Arg	Leu	Leu
							95			100			105	
His	Lys	Gln	Arg	Leu	Leu	Phe	Ser	Cys	Leu	Leu	Trp	Leu	Thr	Phe
							110			115			120	
Met	Tyr	Phe	Trp	Lys	Leu	Gly	Asp	Leu	Phe	Pro	Ile	Leu	Ser	
							125			130			135	
Pro	Lys	His	Gly	Ile	Leu	Ser	Ile	Glu	Gln	Leu	Ile	Ser	Arg	Val
							140			145			150	
Gly	Val	Ile	Gly	Val	Thr	Leu	Met	Ala	Leu	Leu	Ser	Gly	Phe	Gly
							155			160			165	
Ala	Val	Asn	Cys	Pro	Tyr	Thr	Tyr	Met	Ser	Tyr	Phe	Leu	Arg	Asn
							170			175			180	
Val	Thr	Asp	Thr	Asp	Ile	Leu	Ala	Leu	Glu	Arg	Arg	Leu	Leu	Gln
							185			190			195	
Thr	Met	Asp	Met	Ile	Ile	Ser	Lys	Lys	Lys	Arg	Met	Ala	Met	Ala
							200			205			210	
Arg	Arg	Thr	Met	Phe	Gln	Lys	Gly	Glu	Val	His	Asn	Lys	Pro	Ser
							215			220			225	
Gly	Phe	Trp	Gly	Met	Ile	Lys	Ser	Val	Thr	Thr	Ser	Ala	Ser	Gly
							230			235			240	
Ser	Glu	Asn	Leu	Thr	Leu	Ile	Gln	Gln	Glu	Val	Asp	Ala	Leu	Glu
							245			250			255	
Glu	Leu	Ser	Arg	Gln	Leu	Phe	Leu	Glu	Thr	Ala	Asp	Leu	Tyr	Ala
							260			265			270	
Thr	Lys	Glu	Arg	Ile	Glu	Tyr	Ser	Lys	Thr	Phe	Lys	Gly	Lys	Tyr
							275			280			285	
Phe	Asn	Phe	Leu	Gly	Tyr	Phe	Phe	Ser	Ile	Tyr	Cys	Val	Trp	Lys
							290			295			300	
Ile	Phe	Met	Ala	Thr	Ile	Asn	Ile	Val	Phe	Asp	Arg	Val	Gly	Lys
							305			310			315	
Thr	Asp	Pro	Val	Thr	Arg	Gly	Ile	Glu	Ile	Thr	Val	Asn	Tyr	Leu
							320			325			330	
Gly	Ile	Gln	Phe	Asp	Val	Lys	Phe	Trp	Ser	Gln	His	Ile	Ser	Phe
							335			340			345	
Ile	Leu	Val	Gly	Ile	Ile	Ile	Val	Thr	Ser	Ile	Arg	Gly	Leu	Leu
							350			355			360	
Ile	Thr	Leu	Thr	Lys	Phe	Phe	Tyr	Ala	Ile	Ser	Ser	Ser	Lys	Ser
							365			370			375	
Ser	Asn	Val	Ile	Val	Leu	Leu	Leu	Ala	Gln	Ile	Met	Gly	Met	Tyr
							380			385			390	
Phe	Val	Ser	Ser	Val	Leu	Leu	Ile	Arg	Met	Ser	Met	Pro	Leu	Glu
							395			400			405	
Tyr	Arg	Thr	Ile	Ile	Thr	Glu	Val	Leu	Gly	Glu	Leu	Gln	Phe	Asn
							410			415			420	
Phe	Tyr	His	Arg	Trp	Phe	Asp	Val	Ile	Phe	Leu	Val	Ser	Ala	Leu
							425			430			435	
Ser	Ser	Ile	Leu	Phe	Leu	Tyr	Leu	Ala	His	Lys	Gln	Ala	Pro	Glu
							440			445			450	

Lys Gln Met Ala Pro
455

<210> 15

<211> 277

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1334051CD1

<400> 15

Met Lys Ile Ser Met Ile Asn Tyr Lys Ser Leu Leu Ala Leu Leu
1 5 10 15
Phe Ile Leu Ala Ser Trp Ile Ile Phe Thr Val Phe Gln Asn Ser
20 25 30
Thr Lys Val Trp Ser Ala Leu Asn Leu Ser Ile Ser Leu His Tyr
35 40 45
Trp Asn Asn Ser Thr Lys Ser Leu Phe Pro Lys Thr Pro Leu Ile
50 55 60
Ser Leu Lys Pro Leu Thr Glu Thr Glu Leu Arg Ile Lys Glu Ile
65 70 75
Ile Glu Lys Leu Asp Gln Gln Ile Pro Pro Arg Pro Phe Thr His
80 85 90
Val Asn Thr Thr Thr Ser Ala Thr His Ser Thr Ala Thr Ile Leu
95 100 105
Asn Pro Arg Asp Thr Tyr Cys Arg Gly Asp Gln Leu His Ile Leu
110 115 120
Leu Glu Val Arg Asp His Leu Gly Arg Arg Lys Gln Tyr Gly Gly
125 130 135
Asp Phe Leu Arg Ala Arg Met Ser Ser Pro Ala Leu Met Ala Gly
140 145 150
Ala Ser Gly Lys Val Thr Asp Phe Asn Asn Gly Thr Tyr Leu Val
155 160 165
Ser Phe Thr Leu Phe Trp Glu Gly Gln Val Ser Leu Ser Leu Leu
170 175 180
Leu Ile His Pro Ser Glu Gly Val Ser Ala Leu Trp Ser Ala Arg
185 190 195
Asn Gln Gly Tyr Asp Arg Val Ile Phe Thr Gly Gln Phe Val Asn
200 205 210
Gly Thr Ser Gln Val His Ser Glu Cys Gly Leu Ile Leu Asn Thr
215 220 225
Asn Ala Glu Leu Cys Gln Tyr Leu Asp Asn Arg Asp Gln Glu Gly
230 235 240
Phe Tyr Cys Val Arg Pro Gln His Met Pro Cys Ala Ala Leu Thr
245 250 255
His Met Tyr Ser Lys Asn Lys Lys Val Ser Tyr Leu Ser Lys Gln
260 265 270
Glu Lys Ser Leu Phe Glu Arg
275

<210> 16

<211> 647

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1336728CD1

<400> 16

Met Ala Ser Leu Val Ser Leu Glu Leu Gly Leu Leu Ala Val
1 5 10 15
Leu Val Val Thr Ala Thr Ala Ser Pro Pro Ala Gly Leu Leu Ser

20	25	30
Leu Leu Thr Ser Gly Gln Gly Ala Leu Asp Gln Glu Ala Leu Gly		
35	40	45
Gly Leu Leu Asn Thr Leu Ala Asp Arg Val His Cys Thr Asn Gly		
50	55	60
Pro Cys Gly Lys Cys Leu Ser Val Glu Asp Ala Leu Gly Leu Gly		
65	70	75
Glu Pro Glu Gly Ser Gly Leu Pro Pro Gly Pro Val Leu Glu Ala		
80	85	90
Arg Tyr Val Ala Arg Leu Ser Ala Ala Ala Val Leu Tyr Leu Ser		
95	100	105
Asn Pro Glu Gly Thr Cys Glu Asp Thr Arg Ala Gly Leu Trp Ala		
110	115	120
Ser His Ala Asp His Leu Leu Ala Leu Leu Glu Ser Pro Lys Ala		
125	130	135
Leu Thr Pro Gly Leu Ser Trp Leu Leu Gln Arg Met Gln Ala Arg		
140	145	150
Ala Ala Gly Gln Thr Pro Lys Thr Ala Cys Val Asp Ile Pro Gln		
155	160	165
Leu Leu Glu Glu Ala Val Gly Ala Gly Ala Pro Gly Ser Ala Gly		
170	175	180
Gly Val Leu Ala Ala Leu Leu Asp His Val Arg Ser Gly Ser Cys		
185	190	195
Phe His Ala Leu Pro Ser Pro Gln Tyr Phe Val Asp Phe Val Phe		
200	205	210
Gln Gln His Ser Ser Glu Val Pro Met Thr Leu Ala Glu Leu Ser		
215	220	225
Ala Leu Met Gln Arg Leu Gly Val Gly Arg Glu Ala His Ser Asp		
230	235	240
His Ser His Arg His Arg Gly Ala Ser Ser Arg Asp Pro Val Pro		
245	250	255
Leu Ile Ser Ser Ser Asn Ser Ser Ser Val Trp Asp Thr Val Cys		
260	265	270
Leu Ser Ala Arg Asp Val Met Ala Ala Tyr Gly Leu Ser Glu Gln		
275	280	285
Ala Gly Val Thr Pro Glu Ala Trp Ala Gln Leu Ser Pro Ala Leu		
290	295	300
Leu Gln Gln Gln Leu Ser Gly Ala Cys Thr Ser Gln Ser Arg Pro		
305	310	315
Pro Val Gln Asp Gln Leu Ser Gln Ser Glu Arg Tyr Leu Tyr Gly		
320	325	330
Ser Leu Ala Thr Leu Leu Ile Cys Leu Cys Ala Val Phe Gly Leu		
335	340	345
Leu Leu Leu Thr Cys Thr Gly Cys Arg Gly Val Thr His Tyr Ile.		
350	355	360
Leu Gln Thr Phe Leu Ser Leu Ala Val Gly Ala Leu Thr Gly Asp		
365	370	375
Ala Val Leu His Leu Thr Pro Lys Val Leu Gly Leu His Thr His		
380	385	390
Ser Glu Glu Gly Leu Ser Pro Gln Pro Thr Trp Arg Leu Leu Ala		
395	400	405
Met Leu Ala Gly Leu Tyr Ala Phe Phe Leu Phe Glu Asn Leu Phe		
410	415	420
Asn Leu Leu Leu Pro Arg Asp Pro Glu Asp Leu Glu Asp Gly Pro		
425	430	435
Cys Gly His Ser Ser His Ser His Gly Gly His Ser His Gly Val		
440	445	450
Ser Leu Gln Leu Ala Pro Ser Glu Leu Arg Gln Pro Lys Pro Pro		
455	460	465
His Glu Gly Ser Arg Ala Asp Leu Val Ala Glu Glu Ser Pro Glu		
470	475	480
Leu Leu Asn Pro Glu Pro Arg Arg Leu Ser Pro Glu Leu Arg Leu		
485	490	495

Leu Pro Tyr Met Ile Thr Leu Gly Asp Ala Val His Asn Phe Ala
 500 505 510
 Asp Gly Leu Ala Val Gly Ala Ala Phe Ala Ser Ser Trp Lys Thr
 515 520 525
 Gly Leu Ala Thr Ser Leu Ala Val Phe Cys His Glu Leu Pro His
 530 535 540
 Glu Leu Gly Asp Phe Ala Ala Leu Leu His Ala Gly Leu Ser Val
 545 550 555
 Arg Gln Ala Leu Leu Leu Asn Leu Ala Ser Ala Leu Thr Ala Phe
 560 565 570
 Ala Gly Leu Tyr Val Ala Leu Ala Val Gly Val Ser Glu Glu Ser
 575 580 585
 Glu Ala Trp Ile Leu Ala Val Ala Thr Gly Leu Phe Leu Tyr Val
 590 595 600
 Ala Leu Cys Asp Met Leu Pro Ala Met Leu Lys Val Arg Asp Pro
 605 610 615
 Arg Pro Trp Leu Leu Phe Leu Leu His Asn Val Gly Leu Leu Gly
 620 625 630
 Gly Trp Thr Val Leu Leu Leu Ser Leu Tyr Glu Asp Asp Ile
 635 640 645
 Thr Phe

<210> 17
 <211> 406
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1452856CD1

<400> 17
 Met Ala Glu Asn Gly Lys Asn Cys Asp Gln Arg Arg Val Ala Met
 1. 5 10 15
 Asn Lys Glu His His Asn Gly Asn Phe Thr Asp Pro Ser Ser Val
 20 25 30
 Asn Glu Lys Lys Arg Arg Glu Arg Glu Arg Gln Asn Ile Val
 35 40 45
 Leu Trp Arg Gln Pro Leu Ile Thr Leu Gln Tyr Phe Ser Leu Glu
 50 55 60
 Ile Leu Val Ile Leu Lys Glu Trp Thr Ser Lys Leu Trp His Arg
 65 70 75
 Gln Ser Ile Val Val Ser Phe Leu Leu Leu Ala Val Leu Ile
 80 85 90
 Ala Thr Tyr Tyr Val Glu Gly Val His Gln Gln Tyr Val Gln Arg
 95 100 105
 Ile Glu Lys Gln Phe Leu Leu Tyr Ala Tyr Trp Ile Gly Leu Gly
 110 115 120
 Ile Leu Ser Ser Val Gly Leu Gly Thr Gly Leu His Thr Phe Leu
 125 130 135
 Leu Tyr Leu Gly Pro His Ile Ala Ser Val Thr Leu Ala Ala Tyr
 140 145 150
 Glu Cys Asn Ser Val Asn Phe Pro Glu Pro Pro Tyr Pro Asp Gln
 155 160 165
 Ile Ile Cys Pro Asp Glu Glu Gly Thr Glu Gly Thr Ile Ser Leu
 170 175 180
 Trp Ser Ile Ile Ser Lys Val Arg Ile Glu Ala Cys Met Trp Gly
 185 190 195
 Ile Gly Thr Ala Ile Gly Glu Leu Pro Pro Tyr Phe Met Ala Arg
 200 205 210
 Ala Ala Arg Leu Ser Gly Ala Glu Pro Asp Asp Glu Glu Tyr Gln
 215 220 225
 Glu Phe Glu Glu Met Leu Glu His Ala Glu Ser Ala Gln Asp Phe

230	235	240
Ala Ser Arg Ala Lys Leu Ala Val Gln	Lys Leu Val Gln Lys Val	
245	250	255
Gly Phe Phe Gly Ile Leu Ala Cys Ala	Ser Ile Pro Asn Pro Leu	
260	265	270
Phe Asp Leu Ala Gly Ile Thr Cys Gly	His Phe Leu Val Pro Phe	
275	280	285
Trp Thr Phe Phe Gly Ala Thr Leu Ile	Gly Lys Ala Ile Ile Lys	
290	295	300
Met His Ile Gln Lys Ile Phe Val Ile	Ile Thr Phe Ser Lys His	
305	310	315
Ile Val Glu Gln Met Val Ala Phe Ile	Gly Ala Val Pro Gly Ile	
320	325	330
Gly Pro Ser Leu Gln Lys Pro Phe Gln	Glu Tyr Leu Glu Ala Gln	
335	340	345
Arg Gln Lys Leu His His Lys Ser Glu	Met Gly Thr Pro Gln Gly	
350	355	360
Glu Asn Trp Leu Ser Trp Met Phe Glu	Lys Leu Val Val Val Met	
365	370	375
Val Cys Tyr Phe Ile Leu Ser Ile Ile	Asn Ser Met Ala Gln Ser	
380	385	390
Tyr Ala Lys Arg Ile Gln Gln Arg Leu	Asn Ser Glu Glu Lys Thr	
395	400	405
Lys		

<210> 18

<211> 290

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1562471CD1

<400> 18

Met Pro Leu Leu Thr Leu Tyr Leu Leu Leu	Phe Trp Leu Ser Gly		
1	5	10	15
Tyr Ser Ile Ala Thr Gln Ile Thr Gly	Pro Thr Thr Val Asn Gly		
20	25	30	
Leu Glu Arg Gly Ser Leu Thr Val Gln	Cys Val Tyr Arg Ser Gly		
35	40	45	
Trp Glu Thr Tyr Leu Lys Trp Trp Cys	Arg Gly Ala Ile Trp Arg		
50	55	60	
Asp Cys Lys Ile Leu Val Lys Thr Ser	Gly Ser Glu Gln Glu Val		
65	70	75	
Lys Arg Asp Arg Val Ser Ile Lys Asp Asn	Gln Lys Asn Arg Thr		
80	85	90	
Phe Thr Val Thr Met Glu Asp Leu Met	Lys Thr Asp Ala Asp Thr		
95	100	105	
Tyr Trp Cys Gly Ile Glu Lys Thr Gly	Asn Asp Leu Gly Val Thr		
110	115	120	
Val Gln Val Thr Ile Asp Pro Ala Pro	Val Thr Gln Glu Glu Thr		
125	130	135	
Ser Ser Ser Pro Thr Leu Thr Gly His	His Leu Asp Asn Arg His		
140	145	150	
Lys Leu Leu Lys Leu Ser Val Leu Leu	Pro Leu Ile Phe Thr Ile		
155	160	165	
Leu Leu Leu Leu Val Ala Ala Ser	Leu Leu Ala Trp Arg Met		
170	175	180	
Met Lys Tyr Gln Gln Lys Ala Ala Gly	Met Ser Pro Glu Gln Val		
185	190	195	
Leu Gln Pro Leu Glu Gly Asp Leu Cys	Tyr Ala Asp Leu Thr Leu		
200	205	210	

Gln Leu Ala Gly Thr Ser Pro Arg Lys Ala Thr Thr Lys Leu Ser
 215 220 225
 Ser Ala Gln Val Asp Gln Val Glu Val Glu Tyr Val Thr Met Ala
 230 235 240
 Ser Leu Pro Lys Glu Asp Ile Ser Tyr Ala Ser Leu Thr Leu Gly
 245 250 255
 Ala Glu Asp Gln Glu Pro Thr Tyr Cys Asn Met Gly His Leu Ser
 260 265 270
 Ser His Leu Pro Gly Arg Gly Pro Glu Glu Pro Thr Glu Tyr Ser
 275 280 285
 Thr Ile Ser Arg Pro
 290

<210> 19
 <211> 390
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1618158CD1

<400> 19

Met Phe Ser Thr Asn Tyr Ser His Met Glu Asn Tyr Arg Lys Arg
 1 5 10 15
 Glu Asp Leu Val Tyr Gln Ser Thr Val Arg Leu Pro Glu Val Arg
 20 25 30
 Ile Ser Asp Asn Gly Pro Tyr Glu Cys His Val Gly Ile Tyr Asp
 35 40 45
 Arg Ala Thr Arg Glu Lys Val Val Leu Ala Ser Gly Asn Ile Phe
 50 55 60
 Leu Asn Val Met Ala Pro Pro Thr Ser Ile Glu Val Val Ala Ala
 65 70 75
 Asp Thr Pro Ala Pro Phe Ser Arg Tyr Gln Ala Gln Asn Phe Thr
 80 85 90
 Leu Val Cys Ile Val Ser Gly Gly Lys Pro Ala Pro Met Val Tyr
 95 100 105
 Phe Lys Arg Asp Gly Glu Pro Ile Asp Ala Val Pro Leu Ser Glu
 110 115 120
 Pro Pro Ala Ala Ser Ser Gly Pro Leu Gln Asp Ser Arg Pro Phe
 125 130 135
 Arg Ser Leu Leu His Arg Asp Leu Asp Asp Thr Lys Met Gln Lys
 140 145 150
 Ser Leu Ser Leu Leu Asp Ala Glu Asn Arg Gly Gly Arg Pro Tyr
 155 160 165
 Thr Glu Arg Pro Ser Arg Gly Leu Thr Pro Asp Pro Asn Ile Leu
 170 175 180
 Leu Gln Pro Thr Thr Glu Asn Ile Pro Glu Thr Val Val Ser Arg
 185 190 195
 Glu Phe Pro Arg Trp Val His Ser Ala Glu Pro Thr Tyr Phe Leu
 200 205 210
 Arg His Ser Arg Thr Pro Ser Ser Asp Gly Thr Val Glu Val Arg
 215 220 225
 Ala Leu Leu Thr Trp Thr Leu Asn Pro Gln Ile Asp Asn Glu Ala
 230 235 240
 Leu Phe Ser Cys Glu Val Lys His Pro Ala Leu Ser Met Pro Met
 245 250 255
 Gln Ala Glu Val Thr Leu Val Ala Pro Lys Gly Pro Lys Ile Val
 260 265 270
 Met Thr Pro Ser Arg Ala Arg Val Gly Asp Thr Val Arg Ile Leu
 275 280 285
 Val His Gly Phe Gln Asn Glu Val Phe Pro Glu Pro Met Phe Thr
 290 295 300
 Trp Thr Arg Val Gly Ser Arg Leu Leu Asp Gly Ser Ala Glu Phe

305	310	315												
Asp	Gly	Lys	Glu	Leu	Val	Leu	Glu	Arg	Val	Pro	Ala	Glu	Leu	Asn
320														330
Gly	Ser	Met	Tyr	Arg	Cys	Thr	Ala	Gln	Asn	Pro	Leu	Gly	Ser	Thr
335														345
Asp	Thr	His	Thr	Arg	Leu	Ile	Val	Phe	Glu	Asn	Pro	Asn	Ile	Pro
350														360
Arg	Gly	Thr	Glu	Asp	Ser	Asn	Gly	Ser	Ile	Gly	Pro	Thr	Gly	Ala
365														375
Arg	Leu	Thr	Leu	Val	Leu	Ala	Leu	Thr	Val	Ile	Leu	Glu	Leu	Thr
380														390

<210> 20

<211> 427

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1656935CD1

<400> 20

Met	Asn	Val	Asn	Ser	Met	Asp	Met	Thr	Gly	Gly	Leu	Ser	Val	Lys
1									10					15
Asp	Pro	Ser	Gln	Ser	Gln	Ser	Arg	Leu	Pro	Gln	Trp	Thr	His	Pro
														20
														25
Asn	Ser	Met	Asp	Asn	Leu	Pro	Ser	Ala	Ala	Ser	Pro	Leu	Glu	Gln
														35
Asn	Pro	Ser	Lys	His	Gly	Ala	Ile	Pro	Gly	Gly	Leu	Ser	Ile	Gly
														50
Pro	Pro	Gly	Lys	Ser	Ser	Ile	Asp	Asp	Ser	Tyr	Gly	Arg	Tyr	Asp
														65
Leu	Ile	Gln	Asn	Ser	Glu	Ser	Pro	Ala	Ser	Pro	Pro	Val	Ala	Val
														80
Pro	His	Ser	Trp	Ser	Arg	Ala	Lys	Ser	Asp	Ser	Asp	Lys	Ile	Ser
														95
Asn	Gly	Ser	Ser	Ile	Asn	Trp	Pro	Pro	Glu	Phe	His	Pro	Gly	Val
														110
Pro	Trp	Lys	Gly	Leu	Gln	Asn	Ile	Asp	Pro	Glu	Asn	Asp	Pro	Asp
														125
Val	Thr	Pro	Gly	Ser	Val	Pro	Thr	Gly	Pro	Thr	Ile	Asn	Thr	Thr
														140
Ile	Gln	Asp	Val	Asn	Arg	Tyr	Leu	Leu	Lys	Ser	Gly	Gly	Ser	Ser
														155
Pro	Pro	Ser	Ser	Gln	Asn	Ala	Thr	Leu	Pro	Ser	Ser	Ser	Ala	Trp
														170
Pro	Leu	Ser	Ala	Ser	Gly	Tyr	Ser	Ser	Ser	Phe	Ser	Ser	Ile	Ala
														185
Ser	Ala	Pro	Ser	Val	Ala	Gly	Lys	Leu	Ser	Asp	Ile	Lys	Ser	Thr
														200
Trp	Ser	Ser	Gly	Pro	Thr	Ser	His	Thr	Gln	Ala	Ser	Leu	Ser	His
														215
Glu	Leu	Trp	Lys	Val	Pro	Arg	Asn	Ser	Thr	Ala	Pro	Thr	Arg	Pro
														230
Pro	Pro	Gly	Leu	Thr	Asn	Pro	Lys	Pro	Ser	Ser	Thr	Trp	Gly	Ala
														245
Ser	Pro	Leu	Gly	Trp	Thr	Ser	Ser	Tyr	Ser	Ser	Gly	Ser	Ala	Trp
														260
Ser	Thr	Asp	Thr	Ser	Gly	Arg	Thr	Ser	Ser	Trp	Leu	Val	Leu	Arg
														275
Asn	Leu	Thr	Pro	Gln	Ile	Asp	Gly	Ser	Lys	Leu	Arg	Thr	Leu	Cys
														290
Leu	Gln	His	Gly	Pro	Leu	Ile	Thr	Phe	His	Leu	Asn	Leu	Thr	Gln
														295

305	310	315
Gly Asn Ala Val Val Arg Tyr Ser Ser	Lys Glu Glu Gly Leu	Pro
320	325	330
Lys Ala Gln Glu Val Leu Cys Thr Ile	Val Arg Pro Trp Glu	Thr
335	340	345
Leu Ser His Ser Leu Gly Pro Ser Phe	Arg Leu Val Gly Thr	Lys
350	355	360
Glu Val Gly Ile Arg Val Ser Phe Lys	Pro Pro Glu Gly Pro	Gly
365	370	375
Arg Ile Gly Gln Ser Thr Ile Phe Gln	Gly Leu Ala Gln Phe	His
380	385	390
Asp Gln Arg Gly Val Ser Lys Leu Thr	Gly Arg Gly Gly Ile	His
395	400	405
Arg Pro Arg Gly Arg Gly Lys Ala Ser	His Gln Leu Ala His	Met
410	415	420
Arg His Cys Glu Leu Thr Phe		
425		

<210> 21

<211> 459

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1859305CD1

<400> 21

Met Glu Lys Thr Cys Ile Asp Ala Leu Pro	Leu Thr Met Asn Ser	
1 5	10	15
Ser Glu Lys Gln Glu Thr Val Cys Ile	Phe Gly Thr Gly Asp Phe	
20 25	30	
Gly Arg Ser Leu Gly Leu Lys Met Leu	Gln Cys Gly Tyr Ser Val	
35 40	45	
Val Phe Gly Ser Arg Asn Pro Gln Lys	Thr Thr Leu Leu Pro Ser	
50 55	60	
Gly Ala Glu Val Leu Ser Tyr Ser Glu	Ala Ala Lys Lys Ser Asp	
65 70	75	
Ile Ile Ile Ile Ala Ile His Arg Glu	His Tyr Asp Phe Leu Thr	
80 85	90	
Glu Leu Thr Glu Val Leu Asn Gly Lys	Ile Leu Val Asp Ile Ser	
95 100	105	
Asn Asn Leu Lys Ile Asn Gln Tyr Pro	Glu Ser Asn Ala Glu Tyr	
110 115	120	
Leu Ala His Leu Val Pro Gly Ala His	Val Val Lys Ala Phe Asn	
125 130	135	
Thr Ile Ser Ala Trp Ala Leu Gln Ser	Gly Ala Leu Asp Ala Ser	
140 145	150	
Arg Gln Val Phe Val Cys Gly Asn Asp	Ser Lys Ala Lys Gln Arg	
155 160	165	
Val Met Asp Ile Val Arg Asn Leu Gly	Leu Thr Pro Met Asp Gln	
170 175	180	
Gly Ser Leu Met Ala Ala Lys Glu Ile	Glu Lys Tyr Pro Leu Gln	
185 190	195	
Leu Phe Pro Met Trp Arg Phe Pro Phe	Tyr Leu Ser Ala Val Leu	
200 205	210	
Cys Val Phe Leu Phe Phe Tyr Cys Val	Ile Arg Asp Val Ile Tyr	
215 220	225	
Pro Tyr Val Tyr Glu Lys Lys Asp Asn	Thr Phe Arg Met Ala Ile	
230 235	240	
Ser Ile Pro Asn Arg Ile Phe Pro Ile	Leu Ala Leu Thr Leu Leu	
245 250	255	
Ala Leu Val Tyr Leu Pro Gly Val Ile	Ala Ala Ile Leu Gln Leu	
260 265	270	

Tyr Arg Gly Thr Lys Tyr Arg Arg Phe Pro Asp Trp Leu Asp His
 275 280 285
 Trp Met Leu Cys Arg Lys Gln Leu Gly Leu Val Ala Leu Gly Phe
 290 295 300
 Ala Phe Leu His Val Leu Tyr Thr Leu Val Ile Pro Ile Arg Tyr
 305 310 315
 Tyr Val Arg Trp Arg Leu Gly Asn Leu Thr Val Thr Gln Ala Ile
 320 325 330
 Leu Lys Lys Glu Asn Pro Phe Ser Thr Ser Ser Ala Trp Leu Ser
 335 340 345
 Asp Ser Tyr Val Ala Leu Gly Ile Leu Gly Phe Phe Leu Phe Val
 350 355 360
 Leu Leu Gly Ile Thr Ser Leu Pro Ser Val Ser Asn Ala Val Asn
 365 370 375
 Trp Arg Glu Phe Arg Phe Val Gln Ser Lys Leu Gly Tyr Leu Thr
 380 385 390
 Leu Ile Leu Cys Thr Ala His Thr Leu Val Tyr Gly Gly Lys Arg
 395 400 405
 Phe Leu Ser Pro Ser Asn Leu Arg Trp Tyr Leu Pro Ala Ala Tyr
 410 415 420
 Val Leu Gly Leu Ile Ile Pro Cys Thr Val Leu Val Ile Lys Phe
 425 430 435
 Val Leu Ile Met Pro Cys Val Asp Asn Thr Leu Thr Arg Ile Arg
 440 445 450
 Gln Gly Trp Glu Arg Asn Ser Lys His
 455

<210> 22

<211> 229

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1949083CD1

<400> 22

Met Leu Pro Val Ser Arg Thr Cys Leu Leu Glu Ser Ser Thr Arg
 1 5 10 15
 Leu Lys Pro His Glu Ala Gln Asn Tyr Arg Lys Lys Ala Leu Trp
 20 25 30
 Val Ser Trp Phe Ser Ile Ile Val Thr Leu Ala Leu Ala Val Ala
 35 40 45
 Ala Phe Thr Val Ser Val Met Arg Tyr Ser Ala Ser Ala Phe Gly
 50 55 60
 Phe Ala Phe Asp Ala Ile Leu Asp Val Leu Ser Ser Ala Ile Val
 65 70 75
 Leu Trp Arg Tyr Ser Asn Ala Ala Val His Ser Ala His Arg
 80 85 90
 Glu Tyr Ile Ala Cys Val Ile Leu Gly Val Ile Phe Leu Leu Ser
 95 100 105
 Ser Ile Cys Ile Val Val Lys Ala Ile His Asp Leu Ser Thr Arg
 110 115 120
 Leu Leu Pro Glu Val Asp Asp Phe Leu Phe Ser Val Ser Ile Leu
 125 130 135
 Ser Gly Ile Leu Cys Ser Ile Leu Ala Val Leu Lys Phe Met Leu
 140 145 150
 Gly Lys Val Leu Thr Ser Arg Ala Leu Ile Thr Asp Gly Phe Asn
 155 160 165
 Ser Leu Val Gly Gly Val Met Gly Phe Ser Ile Leu Leu Ser Ala
 170 175 180
 Glu Val Phe Lys His Asp Ser Ala Val Trp Tyr Leu Asp Gly Ser
 185 190 195
 Ile Gly Val Leu Ile Gly Leu Thr Ile Phe Ala Tyr Gly Val Lys

200	205	210
Leu Leu Ile Asp Met Val Pro Lys Val Arg Gln Thr Arg His Tyr		
215	220	225

Glu Met Phe Glu

<210> 23

<211> 311

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1996357CD1

<400> 23

Met Ala Val Asp Ile Gln Pro Ala Cys Leu Gly Leu Tyr Cys Gly			
1	5	10	15
Lys Thr Leu Leu Phe Lys Asn Gly Ser Thr Glu Ile Tyr Gly Glu			
20	25		30
Cys Gly Val Cys Pro Arg Gly Gln Arg Thr Asn Ala Gln Lys Tyr			
35	40		45
Cys Gln Pro Cys Thr Glu Ser Pro Glu Leu Tyr Asp Trp Leu Tyr			
50	55		60
Leu Gly Phe Met Ala Met Leu Pro Leu Val Leu His Trp Phe Phe			
65	70		75
Ile Glu Trp Tyr Ser Gly Lys Lys Ser Ser Ser Ala Leu Phe Gln			
80	85		90
His Ile Thr Ala Leu Phe Glu Cys Ser Met Ala Ala Ile Ile Thr			
95	100		105
Leu Leu Val Ser Asp Pro Val Gly Val Leu Tyr Ile Arg Ser Cys			
110	115		120
Arg Val Leu Met Leu Ser Asp Trp Tyr Thr Met Leu Tyr Asn Pro			
125	130		135
Ser Pro Asp Tyr Val Thr Thr Val His Cys Thr His Glu Ala Val			
140	145		150
Tyr Pro Leu Tyr Thr Ile Val Phe Ile Tyr Tyr Ala Phe Cys Leu			
155	160		165
Val Leu Met Met Leu Leu Arg Pro Leu Leu Val Lys Lys Ile Ala			
170	175		180
Cys Gly Leu Gly Lys Ser Asp Arg Phe Lys Ser Ile Tyr Ala Ala			
185	190		195
Leu Tyr Phe Phe Pro Ile Leu Thr Val Leu Gln Ala Val Gly Gly			
200	205		210
Gly Leu Leu Tyr Tyr Ala Phe Pro Tyr Ile Ile Leu Val Leu Ser			
215	220		225
Leu Val Thr Leu Ala Val Tyr Met Ser Ala Ser Glu Ile Glu Asn			
230	235		240
Cys Tyr Asp Leu Leu Val Arg Lys Lys Arg Leu Ile Val Leu Phe			
245	250		255
Ser His Trp Leu Leu His Ala Tyr Gly Ile Ile Ser Ile Ser Arg			
260	265		270
Val Asp Lys Leu Glu Gln Asp Leu Pro Leu Leu Ala Leu Val Pro			
275	280		285
Thr Pro Ala Leu Phe Tyr Leu Phe Thr Ala Lys Phe Thr Glu Pro			
290	295		300
Ser Arg Ile Leu Ser Glu Gly Ala Asn Gly His			
305	310		

<210> 24

<211> 92

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature
 <223> Incyte ID No: 2061330CD1

<400> 24
 Met Arg Phe Ile Phe Leu Lys Phe Trp Thr Tyr Thr Val Arg Ala
 1 5 10 15
 Ser Thr Asp Leu Thr Gln Thr Gly Asp Cys Ser Gln Cys Thr His
 20 25 30
 Gln Val Thr Glu Val Gly Gln Gln Ile Lys Thr Ile Phe Leu Phe
 35 40 45
 Tyr Ser Tyr Tyr Glu Cys Met Glu Thr Ile Lys Glu Thr Cys Leu
 50 55 60
 Tyr Asn Ala Thr Gln Tyr Lys Val Cys Ser Pro Arg Asn Asp Arg
 65 70 75
 Pro Asp Val Cys Tyr Asn Pro Ser Glu Pro Pro Ala Pro Pro Phe
 80 85 90
 Leu Lys

<210> 25
 <211> 258
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2346947CD1

<400> 25
 Met Ala Glu Ser Pro Gly Cys Cys Ser Val Trp Ala Arg Cys Leu
 1 5 10 15
 His Cys Leu Tyr Ser Cys His Trp Arg Lys Cys Pro Arg Glu Arg
 20 25 30
 Met Gln Thr Ser Lys Cys Asp Cys Ile Trp Phe Gly Leu Leu Phe
 35 40 45
 Leu Thr Phe Leu Leu Ser Leu Ser Trp Leu Tyr Ile Gly Leu Val
 50 55 60
 Leu Leu Asn Asp Leu His Asn Phe Asn Glu Phe Leu Phe Arg Arg
 65 70 75
 Trp Gly His Trp Met Asp Trp Ser Leu Ala Phe Leu Leu Val Ile
 80 85 90
 Ser Leu Leu Val Thr Tyr Ala Ser Leu Leu Leu Val Leu Ala Leu
 95 100 105
 Leu Leu Arg Leu Cys Arg Gln Pro Leu His Leu His Ser Leu His
 110 115 120
 Lys Val Leu Leu Leu Leu Ile Met Leu Leu Val Ala Ala Gly Leu
 125 130 135
 Val Gly Leu Asp Ile Gln Trp Gln Gln Glu Trp His Ser Leu Arg
 140 145 150
 Val Ser Leu Gln Ala Thr Ala Pro Phe Leu His Ile Gly Ala Ala
 155 160 165
 Ala Gly Ile Ala Leu Leu Ala Trp Pro Val Ala Asp Thr Phe Tyr
 170 175 180
 Arg Ile His Arg Arg Gly Pro Lys Ile Leu Leu Leu Leu Phe
 185 190 195
 Phe Gly Val Val Leu Val Ile Tyr Leu Ala Pro Leu Cys Ile Ser
 200 205 210
 Ser Pro Cys Ile Met Glu Pro Arg Asp Leu Pro Pro Lys Pro Gly
 215 220 225
 Leu Val Gly His Arg Gly Ala Pro Met Leu Ala Pro Glu Asn Thr
 230 235 240
 Leu Met Ser Leu Arg Lys Thr Ala Glu Cys Gly Leu Leu Cys Leu
 245 250 255
 Arg Leu Met

<210> 26
<211> 226
<212> PRT
<213> *Homo sapiens*

<220>
<221> misc_feature
<223> Incyte ID No: 2795577CD1

<400> 26
 Met Lys Met Val Ala Pro Trp Thr Arg Phe Tyr Ser Asn Ser Cys
 1 5 10 15
 Cys Leu Cys Cys His Val Arg Thr Gly Thr Ile Leu Leu Gly Val
 20 25 30
 Trp Tyr Leu Ile Ile Asn Ala Val Val Leu Leu Ile Leu Leu Ser
 35 40 45
 Ala Leu Ala Asp Pro Asp Gln Tyr Asn Phe Ser Ser Ser Glu Leu
 50 55 60
 Gly Gly Asp Phe Glu Phe Met Asp Asp Ala Asn Met Cys Ile Ala
 65 70 75
 Ile Ala Ile Ser Leu Leu Met Ile Leu Ile Cys Ala Met Ala Thr
 80 85 90
 Tyr Gly Ala Tyr Lys Gln Arg Ala Ala Trp Ile Ile Pro Phe Phe
 95 100 105
 Cys Tyr Gln Ile Phe Asp Phe Ala Leu Asn Met Leu Val Ala Ile
 110 115 120
 Thr Val Leu Ile Tyr Pro Asn Ser Ile Gln Glu Tyr Ile Arg Gln
 125 130 135
 Leu Pro Pro Asn Phe Pro Tyr Arg Asp Asp Val Met Ser Val Asn
 140 145 150
 Pro Thr Cys Leu Val Leu Ile Ile Leu Leu Phe Ile Ser Ile Ile
 155 160 165
 Leu Thr Phe Lys Gly Tyr Leu Ile Ser Cys Val Trp Asn Cys Tyr
 170 175 180
 Arg Tyr Ile Asn Gly Arg Asn Ser Ser Asp Val Leu Val Tyr Val
 185 190 195
 Thr Ser Asn Asp Thr Thr Val Leu Leu Pro Pro Tyr Asp Asp Ala
 200 205 210
 Thr Val Asn Gly Ala Ala Lys Glu Pro Pro Pro Pro Tyr Val Ser
 215 220 225
 Ala

<210> 27
<211> 136
<212> PRT
<213> *Homo sapiens*

<220>
<221> misc_feature
<223> Incyte ID No: 3255825CD1

```

<400> 27
Met Ile Ser Ile Thr Glu Trp Gln Lys Ile Gly Val Gly Ile Thr
      1           5           10          15
Gly Phe Gly Ile Phe Phe Ile Leu Phe Gly Thr Leu Leu Tyr Phe
      20          25          30
Asp Ser Val Leu Leu Ala Phe Gly Asn Leu Leu Phe Leu Thr Gly
      35          40          45
Leu Ser Leu Ile Ile Gly Leu Arg Lys Thr Phe Trp Phe Phe Phe
      50          55          60
Gln Arg His Lys Leu Lys Gly Thr Ser Phe Leu Leu Gly Gly Val
      65          70          75

```

Val Ile Val Leu Leu Arg Trp Pro Leu Leu Gly Met Phe Leu Glu
 80 85 90
 Thr Tyr Gly Phe Ser Leu Phe Lys Gly Phe Phe Pro Val Ala
 95 100 105
 Phe Gly Ser Trp Ala Met Ser Ala Thr Ser Pro Ser Trp Val Arg
 110 115 120
 Cys Ser Gly Asp Phe Lys Ala Leu Ala Arg Trp Ser Glu Lys Gln
 125 130 135

Arg

<210> 28
 <211> 458
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3393430CD1

<400> 28
 Met Ala Trp Ala Ser Arg Leu Gly Leu Leu Ala Leu Leu Leu
 1 5 10 15
 Pro Val Val Gly Ala Ser Thr Pro Gly Thr Val Val Arg Leu Asn
 20 25 30
 Lys Ala Ala Leu Ser Tyr Val Ser Glu Ile Gly Lys Ala Pro Leu
 35 40 45
 Gln Arg Ala Leu Gln Val Thr Val Pro His Phe Leu Asp Trp Ser
 50 55 60
 Gly Glu Ala Leu Gln Pro Thr Arg Ile Arg Ile Leu Asn Val His
 65 70 75
 Val Pro Arg Leu His Leu Lys Phe Ile Ala Gly Phe Gly Val Arg
 80 85 90
 Leu Leu Ala Ala Asn Phe Thr Phe Lys Val Phe Arg Ala Pro
 95 100 105
 Glu Pro Leu Glu Leu Thr Leu Pro Val Glu Leu Leu Ala Asp Thr
 110 115 120
 Arg Val Thr Gln Ser Ser Ile Arg Thr Pro Val Val Ser Ile Ser
 125 130 135
 Ala Cys Ser Leu Phe Ser Gly His Ala Asn Glu Phe Asp Gly Ser
 140 145 150
 Asn Ser Thr Ser His Ala Leu Leu Val Leu Val Gln Lys His Ile
 155 160 165
 Lys Ala Val Leu Ser Asn Lys Leu Cys Leu Ser Ile Ser Asn Leu
 170 175 180
 Val Gln Gly Val Asn Val His Leu Gly Thr Leu Ile Gly Leu Asn
 185 190 195
 Pro Val Gly Pro Glu Ser Gln Ile Arg Tyr Ser Met Val Ser Val
 200 205 210
 Pro Thr Val Thr Ser Asp Tyr Ile Ser Leu Glu Val Asn Ala Val
 215 220 225
 Leu Phe Leu Leu Gly Lys Pro Ile Ile Leu Pro Thr Asp Ala Thr
 230 235 240
 Pro Phe Val Leu Pro Arg His Val Gly Thr Glu Gly Ser Met Ala
 245 250 255
 Thr Val Gly Leu Ser Gln Gln Leu Phe Asp Ser Ala Leu Leu
 260 265 270
 Leu Gln Lys Ala Gly Ala Leu Asn Leu Asp Ile Thr Gly Gln Leu
 275 280 285
 Arg Ser Asp Asp Asn Leu Leu Asn Thr Ser Ala Leu Gly Arg Leu
 290 295 300
 Ile Pro Glu Val Ala Arg Gln Phe Pro Glu Pro Met Pro Val Val
 305 310 315
 Leu Lys Val Arg Leu Gly Ala Thr Pro Val Ala Met Leu His Thr

320	325	330
Asn Asn Ala Thr	Leu Arg Leu Gln Pro	Phe Val Glu Val Leu Ala
335	340	345
Thr Ala Ser Asn	Ser Ala Phe Gln Ser	Leu Phe Ser Leu Asp Val
350	355	360
Val Val Asn Leu	Arg Leu Gln Leu Ser	Val Ser Lys Val Lys Leu
365	370	375
Gln Gly Thr Thr	Ser Val Leu Gly Asp	Val Gln Leu Thr Val Ala
380	385	390
Ser Ser Asn Val	Gly Phe Ile Asp Thr	Asp Gln Val Arg Thr Leu
395	400	405
Met Gly Thr Val	Phe Glu Lys Pro Leu	Leu Asp His Leu Asn Ala
410	415	420
Leu Leu Ala Met	Gly Ile Ala Leu Pro	Gly Val Val Asn Leu His
425	430	435
Tyr Val Ala Pro	Glu Ile Phe Val Tyr	Glu Gly Tyr Val Val Ile
440	445	450
Ser Ser Gly Leu	Phe Tyr Gln Ser	
455		

<210> 29

<211> 368

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3490990CD1

<400> 29

Met Phe Gly Gln Asn Leu Glu Val Gln Leu	Ser Ser Ala Arg Thr	
1 5	10	15
Glu Asn Thr Thr Val Val Trp Lys Ser	Phe His Asp Ser Ile Thr	
20	25	30
Leu Ile Val Leu Ser Ser Glu Val Gly	Ile Ser Glu Leu Arg Leu	
35	40	45
Glu Arg Leu Leu Gln Met Val Phe Gly	Ala Met Val Leu Leu Val	
50	55	60
Gly Leu Glu Glu Leu Thr Asn Ile Arg	Asn Val Glu Arg Leu Lys	
65	70	75
Lys Asp Leu Arg Ala Ser Tyr Cys Leu	Ile Asp Ser Phe Leu Gly	
80	85	90
Asp Ser Glu Leu Ile Gly Asp Leu Thr	Gln Cys Val Asp Cys Val	
95	100	105
Ile Pro Pro Glu Gly Ser Leu Leu Gln	Glu Ala Leu Ser Gly Phe	
110	115	120
Ala Glu Ala Ala Gly Thr Thr Phe Val	Ser Leu Val Val Ser Gly	
125	130	135
Arg Val Val Ala Ala Thr Glu Gly Trp	Trp Arg Leu Gly Thr Pro	
140	145	150
Glu Ala Val Leu Leu Pro Trp Leu Val	Gly Ser Leu Pro Pro Gln	
155	160	165
Thr Ala Arg Asp Tyr Pro Val Tyr Leu	Pro His Gly Ser Pro Thr	
170	175	180
Val Pro His Arg Leu Leu Thr Leu Thr	Leu Leu Pro Ser Leu Glu	
185	190	195
Leu Cys Leu Leu Cys Gly Pro Ser Pro	Pro Leu Ser Gln Leu Tyr	
200	205	210
Pro Gln Leu Leu Glu Arg Trp Trp Gln	Pro Leu Leu Asp Pro Leu	
215	220	225
Arg Ala Cys Leu Pro Leu Gly Pro Arg	Ala Leu Pro Ser Gly Phe	
230	235	240
Pro Leu His Thr Asp Ile Leu Gly Leu	Leu Leu Leu His Leu Glu	
245	250	255

<210> 30
<211> 91
<212> PRT
<213> *Homo sapiens*

<220>
<221> misc_feature
<223> Incyte ID No: 3635154CD1

<400> 30

<210> 31
<211> 295
<212> PRT
<213> *Homo sapiens*

<220>
<221> misc_feature
<223> Incyte ID No: 4374347CD1

<400> 31

Met	Gly	Pro	Pro	Ser	Ala	Cys	Pro	His	Arg	Glu	Cys	Ile	Pro	Trp
1									10					15
Gln	Gly	Leu	Leu	Leu	Thr	Ala	Ser	Leu	Leu	Thr	Phe	Trp	Asn	Ala
									20					30
Pro	Thr	Thr	Ala	Trp	Leu	Phe	Ile	Ala	Ser	Ala	Pro	Phe	Glu	Val
									35					45
Ala	Glu	Gly	Glu	Asn	Val	His	Leu	Ser	Val	Val	Tyr	Leu	Pro	Glu
									50					60
Asn	Leu	Tyr	Ser	Tyr	Gly	Trp	Tyr	Lys	Gly	Lys	Thr	Val	Glu	Pro
									65					75
Asn	Gln	Leu	Ile	Ala	Ala	Tyr	Val	Ile	Asp	Thr	His	Val	Arg	Thr
									80					90
Pro	Gly	Pro	Ala	Tyr	Ser	Gly	Arg	Glu	Thr	Ile	Ser	Pro	Ser	Gly

95	100	105
Asp Leu His Phe Gln Asn Val Thr Leu	Glu Asp Thr Gly Tyr	Tyr
110	115	120
Asn Leu Gln Val Thr Tyr Arg Asn Ser	Gln Ile Glu Gln Ala	Ser
125	130	135
His His Leu Arg Val Tyr Glu Ser Val	Ala Gln Pro Ser Ile	Gln
140	145	150
Ala Ser Ser Thr Thr Val Thr Glu Lys	Gly Ser Val Val Leu	Thr
155	160	165
Cys His Thr Asn Asn Thr Gly Thr Ser	Phe Gln Trp Ile Phe	Asn
170	175	180
Asn Gln Arg Leu Gln Val Thr Lys Arg	Met Lys Leu Ser Trp	Phe
185	190	195
Asn His Val Leu Thr Ile Asp Pro Ile	Arg Gln Glu Asp Ala	Gly
200	205	210
Glu Tyr Gln Cys Glu Val Ser Asn Pro	Val Ser Ser Asn Arg	Ser
215	220	225
Asp Pro Leu Lys Leu Thr Val Lys Tyr	Asp Asn Thr Leu Gly	Ile
230	235	240
Leu Ile Gly Val Leu Val Gly Ser Leu	Leu Val Ala Ala Leu	Val
245	250	255
Cys Phe Leu Leu Leu Arg Lys Thr Gly	Arg Ala Ser Asp Gln	Ser
260	265	270
Asp Phe Arg Glu Gln Gln Pro Pro Ala	Ser Thr Pro Gly His	Gly
275	280	285
Pro Ser Asp Ser Ser Asp Ser Ser Ile	Ser	
290	295	

<210> 32

<211> 724

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4596747CD1

<400> 32

Met Phe Asp Thr Thr Pro His Ser Gly Arg	Ser Thr Pro Ser Ser	
1 5	10	15
Ser Pro Ser Leu Arg Lys Arg Leu Gln	Leu Leu Pro Pro Ser Arg	
20	25	30
Pro Pro Pro Glu Pro Glu Pro Gly Thr	Met Val Glu Lys Gly	Ser
35	40	45
Asp Ser Ser Ser Glu Lys Gly Gly Val	Pro Gly Thr Pro Ser Thr	
50	55	60
Gln Ser Leu Gly Ser Arg Asn Phe Ile	Arg Asn Ser Lys Lys	Met
65	70	75
Gln Ser Trp Tyr Ser Met Leu Ser Pro	Thr Tyr Lys Gln Arg Asn	
80	85	90
Glu Asp Phe Arg Lys Leu Phe Ser Lys	Leu Pro Glu Ala Glu Arg	
95	100	105
Leu Ile Val Asp Tyr Ser Cys Ala Leu	Gln Arg Glu Ile Leu	Leu
110	115	120
Gln Gly Arg Leu Tyr Leu Ser Glu Asn	Trp Ile Cys Phe Tyr	Ser
125	130	135
Asn Ile Phe Arg Trp Glu Thr Thr Ile	Ser Ile Gln Leu Lys	Glu
140	145	150
Val Thr Cys Leu Lys Lys Glu Lys Thr	Ala Lys Leu Ile Pro	Asn
155	160	165
Ala Ile Gln Ile Cys Thr Glu Ser Glu	Lys His Phe Phe Thr	Ser
170	175	180
Phe Gly Ala Arg Asp Arg Cys Phe Leu	Leu Ile Phe Arg Leu	Trp
185	190	195

Gln Asn Ala Leu Leu Glu Lys Thr Leu Ser Pro Arg Glu Leu Trp
 200 205 210
 His Leu Val His Gln Cys Tyr Gly Ser Glu Leu Gly Leu Thr Ser
 215 220 225
 Glu Asp Glu Asp Tyr Val Ser Pro Leu Gln Leu Asn Gly Leu Gly
 230 235 240
 Thr Pro Lys Glu Val Gly Asp Val Ile Ala Leu Ser Asp Ile Thr
 245 250 255
 Ser Ser Gly Ala Ala Asp Arg Ser Gln Glu Pro Ser Pro Val Gly
 260 265 270
 Ser Arg Arg Gly His Val Thr Pro Asn Leu Ser Arg Ala Ser Ser
 275 280 285
 Asp Ala Asp His Gly Ala Glu Glu Asp Lys Glu Glu Gln Val Asp
 290 295 300
 Ser Gln Pro Asp Ala Ser Ser Ser Gln Thr Val Thr Pro Val Ala
 305 310 315
 Glu Pro Pro Ser Thr Glu Pro Thr Gln Pro Asp Gly Pro Thr Thr
 320 325 330
 Leu Gly Pro Leu Asp Leu Leu Pro Ser Glu Glu Leu Leu Thr Asp
 335 340 345
 Thr Ser Asn Ser Ser Ser Thr Gly Glu Glu Ala Asp Leu Ala
 350 355 360
 Ala Leu Leu Pro Asp Leu Ser Gly Arg Leu Leu Ile Asn Ser Val
 365 370 375
 Phe His Val Gly Ala Glu Arg Leu Gln Gln Met Leu Phe Ser Asp
 380 385 390
 Ser Pro Phe Leu Gln Gly Phe Leu Gln Gln Cys Lys Phe Thr Asp
 395 400 405
 Val Thr Leu Ser Pro Trp Ser Gly Asp Ser Lys Cys His Gln Arg
 410 415 420
 Arg Val Leu Thr Tyr Thr Ile Pro Ile Ser Asn Pro Leu Gly Pro
 425 430 435
 Lys Ser Ala Ser Val Val Glu Thr Gln Thr Leu Phe Arg Arg Gly
 440 445 450
 Pro Gln Ala Gly Gly Cys Val Val Asp Ser Glu Val Leu Thr Gln
 455 460 465
 Gly Ile Pro Tyr Gln Asp Tyr Phe Tyr Thr Ala His Arg Tyr Cys
 470 475 480
 Ile Leu Gly Leu Ala Arg Asn Lys Ala Arg Leu Arg Val Ser Ser
 485 490 495
 Glu Ile Arg Tyr Arg Lys Gln Pro Trp Ser Leu Val Lys Ser Leu
 500 505 510
 Ile Glu Lys Asn Ser Trp Ser Gly Ile Glu Asp Tyr Phe His His
 515 520 525
 Leu Glu Arg Glu Leu Ala Lys Ala Glu Lys Leu Ser Leu Glu Glu
 530 535 540
 Gly Gly Lys Asp Ala Arg Gly Leu Leu Ser Gly Leu Arg Arg Arg
 545 550 555
 Lys Arg Pro Leu Ser Trp Arg Ala His Gly Asp Gly Pro Gln His
 560 565 570
 Pro Asp Pro Asp Pro Cys Ala Arg Ala Gly Ile His Thr Ser Gly
 575 580 585
 Ser Leu Ser Ser Arg Phe Ser Glu Pro Ser Val Asp Gln Gly Pro
 590 595 600
 Gly Ala Gly Ile Pro Ser Ala Leu Val Leu Ile Ser Ile Val Ile
 605 610 615
 Cys Val Ser Leu Ile Ile Leu Ile Ala Leu Asn Val Leu Leu Phe
 620 625 630
 Tyr Arg Leu Trp Ser Leu Glu Arg Thr Ala His Thr Phe Glu Ser
 635 640 645
 Trp His Ser Leu Ala Leu Ala Lys Gly Lys Phe Pro Gln Thr Ala
 650 655 660
 Thr Glu Trp Ala Glu Ile Leu Ala Leu Gln Lys Gln Phe His Ser

665	670	675
Val Glu Val His Lys Trp Arg Gln Ile	Leu Arg Ala Ser Val	Glu
680	685	690
Leu Leu Asp Glu Met Lys Phe Ser Leu	Glu Lys Leu His Gln	Gly
695	700	705
Ile Thr Val Ser Asp Pro Pro Phe Asp	Thr Gln Pro Arg Pro Asp	
710	715	720

Asp Ser Phe Ser

<210> 33

<211> 331

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5052680CD1

<400> 33

Met Arg Pro Ala Leu Ala Val Gly	Leu Val Phe Ala Gly Cys Cys	
1 5	10	15
Ser Asn Val Ile Phe Leu Glu Leu	Leu Ala Arg Lys His Pro Gly	
20	25	30
Cys Gly Asn Ile Val Thr Phe Ala	Gln Phe Leu Phe Ile Ala Val	
35	40	45
Glu Gly Phe Leu Phe Glu Ala Asp	Leu Gly Arg Lys Pro Pro Ala	
50	55	60
Ile Pro Ile Arg Tyr Tyr Ala Ile	Met Val Thr Met Phe Phe Thr	
65	70	75
Val Ser Val Val Asn Asn Tyr Ala	Leu Asn Leu Asn Ile Ala Met	
80	85	90
Pro Leu His Met Ile Phe Arg Ser	Gly Ser Leu Ile Ala Asn Met	
95	100	105
Ile Leu Gly Ile Ile Leu Lys Lys	Arg Tyr Ser Ile Phe Lys	
110	115	120
Tyr Thr Ser Ile Ala Leu Val Ser	Val Gly Ile Phe Ile Cys Thr	
125	130	135
Phe Met Ser Ala Lys Gln Val Thr	Ser Gln Ser Ser Leu Ser Glu	
140	145	150
Asn Asp Gly Phe Gln Ala Phe Val	Trp Trp Leu Leu Gly Ile Gly	
155	160	165
Ala Leu Thr Phe Ala Leu Leu Met	Ser Ala Arg Met Gly Ile Phe	
170	175	180
Gln Glu Thr Leu Tyr Lys Arg Phe	Gly Lys His Ser Lys Glu Ala	
185	190	195
Leu Phe Tyr Asn His Ala Leu Pro	Leu Pro Gly Phe Val Phe Leu	
200	205	210
Ala Ser Asp Ile Tyr Asp His Ala	Val Leu Phe Asn Lys Ser Glu	
215	220	225
Leu Tyr Glu Ile Pro Val Ile Gly	Val Thr Leu Pro Ile Met Trp	
230	235	240
Phe Tyr Leu Leu Met Asn Ile Ile	Thr Gln Tyr Val Cys Ile Arg	
245	250	255
Gly Val Phe Ile Leu Thr Thr Glu	Cys Ala Ser Leu Thr Val Thr	
260	265	270
Leu Val Val Thr Leu Arg Lys Phe	Val Ser Leu Ile Phe Ser Ile	
275	280	285
Leu Tyr Phe Gln Asn Pro Phe Thr	Leu Leu Trp His Trp Leu Gly	
290	295	300
Leu Phe Val Phe Ile Gly Thr Leu	Met Tyr Thr Glu Val Trp Asn	
305	310	315
Asn Leu Gly Thr Thr Lys Ser Glu	Pro Gln Lys Asp Ser Lys Lys	
320	325	330

Asn

<210> 34
 <211> 398
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5373575CD1

<400> 34
 Met Leu Gly Arg Ser Gly Tyr Arg Ala Leu Pro Leu Gly Asp Phe
 1 5 10 15
 Asp Arg Phe Gln Gln Ser Ser Phe Gly Phe Leu Gly Ser Gln Lys
 20 25 30
 Gly Cys Leu Ser Pro Glu Arg Gly Gly Val Gly Thr Gly Ala Asp
 35 40 45
 Val Pro Gln Ser Trp Pro Ser Cys Leu Cys His Gly Leu Ile Ser
 50 55 60
 Phe Leu Gly Phe Leu Leu Leu Val Thr Phe Pro Ile Ser Gly
 65 70 75
 Trp Phe Ala Leu Lys Ile Val Pro Thr Tyr Glu Arg Met Ile Val
 80 85 90
 Phe Arg Leu Gly Arg Ile Arg Thr Pro Gln Gly Pro Gly Met Val
 95 100 105
 Leu Leu Leu Pro Phe Ile Asp Ser Phe Gln Arg Val Asp Leu Arg
 110 115 120
 Thr Arg Ala Phe Asn Val Pro Pro Cys Lys Leu Ala Ser Lys Asp
 125 130 135
 Gly Ala Val Leu Ser Val Gly Ala Asp Val Gln Phe Arg Ile Trp
 140 145 150
 Asp Pro Val Leu Ser Val Met Thr Val Lys Asp Leu Asn Thr Ala
 155 160 165
 Thr Arg Met Thr Ala Gln Asn Ala Met Thr Lys Ala Leu Leu Lys
 170 175 180
 Arg Pro Leu Arg Glu Ile Gln Met Glu Lys Leu Lys Ile Ser Asp
 185 190 195
 Gln Leu Leu Leu Glu Ile Asn Asp Val Thr Arg Ala Trp Gly Leu
 200 205 210
 Glu Val Asp Arg Val Glu Leu Ala Val Glu Ala Val Leu Gln Pro
 215 220 225
 Pro Gln Asp Ser Pro Ala Gly Pro Asn Leu Asp Ser Thr Leu Gln
 230 235 240
 Gln Leu Ala Leu His Phe Leu Gly Gly Ser Met Asn Ser Met Ala
 245 250 255
 Gly Gly Ala Pro Ser Pro Gly Pro Ala Asp Thr Val Glu Met Val
 260 265 270
 Ser Glu Val Glu Pro Pro Ala Pro Gln Val Gly Ala Arg Ser Ser
 275 280 285
 Pro Lys Gln Pro Leu Ala Glu Gly Leu Leu Thr Ala Leu Gln Pro
 290 295 300
 Phe Leu Ser Glu Ala Leu Val Ser Gln Val Gly Ala Cys Tyr Gln
 305 310 315
 Phe Asn Val Val Leu Pro Ser Gly Thr Gln Ser Ala Tyr Phe Leu
 320 325 330
 Asp Leu Thr Thr Gly Arg Gly Arg Val Gly His Gly Val Pro Asp
 335 340 345
 Gly Ile Pro Asp Val Val Val Glu Met Ala Glu Ala Asp Leu Arg
 350 355 360
 Ala Leu Leu Cys Arg Glu Leu Arg Pro Leu Gly Ala Tyr Met Ser
 365 370 375
 Gly Arg Leu Lys Val Lys Gly Asp Leu Ala Met Ala Met Lys Leu

380	385	390
Glu Ala Val Leu Arg Ala Leu Lys		
395		
<210> 35		
<211> 220		
<212> PRT		
<213> Homo sapiens		
<220>		
<221> misc_feature		
<223> Incyte ID No: 5524468CD1		
<400> 35		
Met Thr Trp Leu Val Leu Leu Gly Thr Leu Leu Cys Met Leu Arg		
1	5	10
Val Gly Leu Gly Thr Pro Asp Ser Glu Gly Phe Pro Pro Arg Ala		
20	25	30
Leu His Asn Cys Pro Tyr Lys Cys Ile Cys Ala Ala Asp Leu Leu		
35	40	45
Ser Cys Thr Gly Leu Gly Leu Gln Asp Val Pro Ala Glu Leu Pro		
50	55	60
Ala Ala Thr Ala Asp Leu Asp Leu Ser His Asn Ala Leu Gln Arg		
65	70	75
Leu Arg Pro Gly Trp Leu Ala Pro Leu Phe Gln Leu Arg Ala Leu		
80	85	90
His Leu Asp His Asn Glu Leu Asp Ala Leu Gly Arg Gly Val Phe		
95	100	105
Val Asn Ala Ser Gly Leu Arg Leu Leu Asp Leu Ser Ser Asn Thr		
110	115	120
Leu Arg Ala Leu Gly Arg His Asp Leu Asp Gly Leu Gly Ala Leu		
125	130	135
Glu Lys Leu Leu Leu Phe Asn Asn Arg Leu Val His Leu Asp Glu		
140	145	150
His Ala Phe His Gly Leu Arg Ala Leu Ser His Leu Tyr Leu Gly		
155	160	165
Cys Asn Glu Leu Ala Ser Phe Ser Phe Asp His Leu His Gly Leu		
170	175	180
Ser Ala Thr His Leu Leu Thr Leu Asp Leu Ser Ser Asn Arg Leu		
185	190	195
Gly His Ile Ser Val Pro Glu Leu Ala Ala Leu Pro Ala Phe Leu		
200	205	210
Lys Asn Gly Leu Tyr Leu His Asp Asn Thr		
215	220	
<210> 36		
<211> 706		
<212> PRT		
<213> Homo sapiens		
<220>		
<221> misc_feature		
<223> Incyte ID No: 5944279CD1		
<400> 36		
Met Glu Glu Asn Pro Thr Leu Glu Ser Glu Ala Trp Gly Ser Ser		
1	5	10
Arg Gly Trp Leu Ala Pro Arg Glu Ala Arg Gly Gly Pro Ser Leu		
20	25	30
Ser Ser Val Leu Asn Glu Leu Pro Ser Ala Ala Thr Leu Arg Tyr		
35	40	45
Arg Asp Pro Gly Val Leu Pro Trp Gly Ala Leu Glu Glu Glu		
50	55	60
Glu Asp Gly Gly Arg Ser Arg Lys Ala Phe Thr Glu Val Thr Gln		
65	70	75

Thr Glu Leu Gln Asp Pro His Pro Ser Arg Glu Leu Pro Trp Pro
 80 85 90
 Met Gln Ala Arg Arg Ala His Arg Gln Arg Asn Ala Ser Arg Asp
 95 100 105
 Gln Val Val Tyr Gly Ser Gly Thr Lys Thr Asp Arg Trp Ala Arg
 110 115 120
 Leu Leu Arg Arg Ser Lys Glu Lys Thr Lys Glu Gly Leu Arg Ser
 125 130 135
 Leu Gln Pro Trp Ala Trp Thr Leu Lys Arg Ile Gly Gly Gln Phe
 140 145 150
 Gly Ala Gly Thr Glu Ser Tyr Phe Ser Leu Leu Arg Phe Leu Leu
 155 160 165
 Leu Leu Asn Val Leu Ala Ser Val Leu Met Ala Cys Met Thr Leu
 170 175 180
 Leu Pro Thr Trp Leu Gly Gly Ala Pro Pro Gly Pro Pro Gly Pro
 185 190 195
 Asp Ile Ser Ser Pro Cys Gly Ser Tyr Asn Pro His Ser Gln Gly
 200 205 210
 Leu Val Thr Phe Ala Thr Gln Leu Phe Asn Leu Leu Ser Gly Glu
 215 220 225
 Gly Tyr Leu Glu Trp Ser Pro Leu Phe Tyr Gly Phe Tyr Pro Pro
 230 235 240
 Arg Pro Arg Leu Ala Val Thr Tyr Leu Cys Trp Ala Phe Ala Val
 245 250 255
 Gly Leu Ile Cys Leu Leu Leu Ile Leu His Arg Ser Val Ser Gly
 260 265 270
 Leu Lys Gln Thr Leu Leu Ala Glu Ser Glu Ala Leu Thr Ser Tyr
 275 280 285
 Ser His Arg Val Phe Ser Ala Trp Asp Phe Gly Leu Cys Gly Asp
 290 295 300
 Val His Val Arg Leu Arg Gln Arg Ile Ile Leu Tyr Glu Leu Lys
 305 310 315
 Val Glu Leu Glu Glu Thr Val Val Arg Arg Gln Ala Ala Val Arg
 320 325 330
 Thr Leu Gly Gln Gln Ala Arg Val Trp Leu Val Arg Val Leu Leu
 335 340 345
 Asn Leu Leu Val Val Ala Leu Leu Gly Ala Ala Phe Tyr Gly Val
 350 355 360
 Tyr Trp Ala Thr Gly Cys Thr Val Glu Leu Gln Glu Met Pro Leu
 365 370 375
 Val Gln Glu Leu Pro Leu Leu Lys Leu Gly Val Asn Tyr Leu Pro
 380 385 390
 Ser Ile Phe Ile Ala Gly Val Asn Phe Val Leu Pro Pro Val Phe
 395 400 405
 Lys Leu Ile Ala Pro Leu Glu Gly Tyr Thr Arg Ser Arg Gln Ile
 410 415 420
 Val Phe Ile Leu Leu Arg Thr Val Phe Leu Arg Leu Ala Ser Leu
 425 430 435
 Val Val Leu Leu Phe Ser Leu Trp Asn Gln Ile Thr Cys Gly Glu
 440 445 450
 Asp Ser Glu Ala Glu Asp Cys Lys Thr Cys Gly Tyr Asn Tyr Lys
 455 460 465
 Gln Leu Pro Cys Trp Glu Thr Val Leu Gly Gln Glu Met Tyr Lys
 470 475 480
 Leu Leu Leu Phe Asp Leu Leu Thr Val Leu Ala Val Ala Leu Leu
 485 490 495
 Ile Gln Phe Pro Arg Lys Leu Leu Cys Gly Leu Cys Pro Gly Ala
 500 505 510
 Leu Gly Arg Leu Ala Gly Thr Gln Glu Phe Gln Val Pro Asp Glu
 515 520 525
 Val Leu Gly Leu Ile Tyr Ala Gln Thr Val Val Trp Val Gly Ser
 530 535 540
 Phe Phe Cys Pro Leu Leu Pro Leu Leu Asn Thr Val Lys Phe Leu

545	550	555
Leu Leu Phe Tyr Leu Lys Lys Leu Thr	Leu Phe Ser Thr Cys	Ser
560	565	570
Pro Ala Ala Arg Thr Phe Arg Ala Ser	Ala Ala Asn Phe Phe	Phe
575	580	585
Pro Leu Val Leu Leu Leu Gly Leu Ala	Ile Ser Ser Val Pro	Leu
590	595	600
Leu Tyr Ser Ile Phe Leu Ile Pro Pro	Ser Lys Leu Cys Gly	Pro
605	610	615
Phe Arg Gly Gln Ser Ser Ile Trp Ala	Gln Ile Pro Glu Ser	Ile
620	625	630
Ser Ser Leu Pro Glu Thr Thr Gln Asn	Phe Leu Phe Phe Leu	Gly
635	640	645
Thr Gln Ala Phe Ala Val Pro Leu Leu	Leu Ile Ser Ser Ile	Leu
650	655	660
Met Ala Tyr Thr Val Ala Leu Ala Asn	Ser Tyr Gly Arg Leu	Ile
665	670	675
Ser Glu Leu Lys Arg Gln Arg Gln Thr	Glu Ala Gln Asn Lys	Val
680	685	690
Phe Leu Ala Arg Arg Ala Val Ala Leu	Thr Ser Thr Lys Pro	Ala
695	700	705

Leu

<210> 37

<211> 466

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6114480CD1

<400> 37

Met Ala Phe Val Leu Ile Leu Val Leu Ser	Phe Tyr Glu Leu Val	
1 5	10	15
Ser Gly Gln Trp Gln Val Val Thr Gly Pro	Gly Lys Phe Val Gln Ala	
20	25	30
Leu Val Gly Glu Asp Ala Val Phe Ser	Cys Ser Leu Phe Pro Glu	
35	40	45
Thr Ser Ala Glu Ala Met Glu Val Arg	Phe Phe Arg Asn Gln Phe	
50	55	60
His Ala Val Val His Leu Tyr Arg Asp	Gly Glu Asp Trp Glu Ser	
65	70	75
Lys Gln Met Pro Gln Tyr Arg Gly Arg	Thr Glu Phe Val Lys Asp	
80	85	90
Ser Ile Ala Gly Gly Arg Val Ser Leu	Arg Leu Lys Asn Ile Thr	
95	100	105
Pro Ser Asp Ile Gly Leu Tyr Gly Cys	Trp Phe Ser Ser Gln Ile	
110	115	120
Tyr Asp Glu Glu Ala Thr Trp Glu Leu	Arg Val Ala Ala Leu Gly	
125	130	135
Ser Leu Pro Leu Ile Ser Ile Val Gly	Tyr Val Asp Gly Gly Ile	
140	145	150
Gln Leu Leu Cys Leu Ser Ser Gly Trp	Phe Pro Gln Pro Thr Ala	
155	160	165
Lys Trp Lys Gly Pro Gln Gly Gln Asp	Leu Ser Ser Asp Ser Arg	
170	175	180
Ala Asn Ala Asp Gly Tyr Ser Leu Tyr	Asp Val Glu Ile Ser Ile	
185	190	195
Ile Val Gln Glu Asn Ala Gly Ser Ile	Leu Cys Ser Ile His Leu	
200	205	210
Ala Glu Gln Ser His Glu Val Glu Ser	Lys Val Leu Ile Gly Glu	
215	220	225

Thr Phe Phe Gln Pro Ser Pro Trp Arg Leu Ala Ser Ile Leu Leu
 230 235 240
 Gly Leu Leu Cys Gly Ala Leu Cys Gly Val Val Met Gly Met Ile
 245 250 255
 Ile Val Phe Phe Lys Ser Lys Gly Lys Ile Gln Ala Glu Leu Asp
 260 265 270
 Trp Arg Arg Lys His Gly Gln Ala Glu Leu Arg Asp Ala Arg Lys
 275 280 285
 His Ala Val Glu Val Thr Leu Asp Pro Glu Thr Ala His Pro Lys
 290 295 300
 Leu Cys Val Ser Asp Leu Lys Thr Val Thr His Arg Lys Ala Pro
 305 310 315
 Gln Glu Val Pro His Ser Glu Lys Arg Phe Thr Arg Lys Ser Val
 320 325 330
 Val Ala Ser Gln Gly Phe Gln Ala Gly Arg His Tyr Trp Glu Val
 335 340 345
 Asp Val Gly Gln Asn Val Gly Trp Tyr Val Gly Val Cys Arg Asp
 350 355 360
 Asp Val Asp Arg Gly Lys Asn Asn Val Thr Leu Ser Pro Asn Asn
 365 370 375
 Gly Tyr Trp Val Leu Arg Leu Thr Thr Glu His Leu Tyr Phe Thr
 380 385 390
 Phe Asn Pro His Phe Ile Ser Leu Pro Pro Ser Thr Pro Pro Thr
 395 400 405
 Arg Val Gly Val Phe Leu Asp Tyr Glu Gly Gly Thr Ile Ser Phe
 410 415 420
 Phe Asn Thr Asn Asp Gln Ser Leu Ile Tyr Thr Leu Leu Thr Cys
 425 430 435
 Gln Phe Glu Gly Leu Leu Arg Pro Tyr Ile Gln His Ala Met Tyr
 440 445 450
 Asp Glu Glu Lys Gly Thr Pro Ile Phe Ile Cys Pro Val Ser Trp
 455 460 465
 Gly

<210> 38
 <211> 2801
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 112301CB1

<400> 38
 cgccttcccc gagcgagacc aaaacagggtg gaatccgggc tggagccgga gctccggcgg 60
 cgcgggtggc ggcacgtccc tccagacagt accacaggca cctggagtag cggcatcggt 120
 cgctgtggcc cccgagtgcc cgtcagagcc taggggagcc tgcctcccg cgcctcgtcg 180
 gggcccgcc aggcaccttg gccgcccgcg cacggacgcg ggcacgagca ctagatcacg 240
 gctgtggac ctccgcacgt tgacaagatt tctctgggtt accgcggaggg attactttga 300
 atttcggtgg tcggctgtgg tctggcatat ttagaactta agtctattat ttcgggcacc 360
 atgactttga ggctttaga agactggtgc agggggatgg acatgaaccc tcggaaagcgc 420
 ctattgattt cccgcatttc ccagagctgc agtgtggcag aaatcgagga ggctctgcag 480
 gctggtttag ctcccttggg ggagtacaga ctgcttggaa ggatgttcag gaggatgtag 540
 aacaggaaag tagccttagt agggcttact gcgagacta gtcacgcctt ggtccctaag 600
 gagataccgg gaaaaggggg tatctggaa gtgatcttta agccccctga cccagataat 660
 acattttaa gcagattaa tgaattttta gcggagagg gcatgacagt gggtagttg 720
 agcagagctc ttggacatga aaatggctcc ttagacccag agcaggccat gatcccgaa 780
 atgtgggccc ctatgttggc acaggcatttta gaggtcttc agcctgcctt gcaatgttt 840
 aagtataaaa agctgagagt gttctcgccg agggagtctc cagaacccagg agaagaagaa 900
 tttggacgct ggatgtttca tactacttag atgataaagg cgtggcagggt gccagatgt 960
 gagaagagaa ggcgattgct agagagcctt cgaggcccag cacttgcgtt tattcgtgtc 1020
 ctcaagataa acaatcctt aattactgtc gatgaatgtc tgcaggctt tgaggaggt 1080
 ttggggta cagataatcc tagggagttt caggtcaaat atctaaccac ttaccagaag 1140

gatggggaaa	agttgtcgcc	ttatgtacta	aggctggagc	ctttgttaca	gaagctggta	1200
cagagaggag	caattgagag	agatgctgt	aatcaggccc	gcctagacca	agtcatgtct	1260
ggggcagttc	acaaaacaat	tcgcagagag	cttaatctgc	cagaggatgg	cccagccct	1320
ggtttcttgc	agttatttgg	actaataaaag	gattatgagg	cagctgagga	ggaggaggcc	1380
cttctccagg	caatatttgg	aggttaattt	acctgagttc	cagggAACCA	cgaagggata	1440
tggcaatgg	tagagcatga	aggtagaaca	gtctatatac	tcttgtgaca	cataacaatcc	1500
ctaccttgc	ctgccaagta	actcattttt	gtcaatttct	cagtataaagc	ccttgcgt	1560
ttctgtgcct	atttaaagtc	tcctaaaggt	gtaatttact	aggaaggatg	tagttctaca	1620
ctgcattta	ccttattttaa	ttcatccctt	tgaatatctt	tgttgggtt	gttggagacag	1680
agtctcgctc	tgtccacccag	gctggagttc	agtgcacgt	tcttgcgtca	ctgcacatctc	1740
cgccttccag	gtttaagtc	ttctctgtcc	tcagttggcc	gtagtgcgtt	gactacagga	1800
atgactcacc	acacccagct	aagttttgc	ttatcagtag	agacgggggtt	tcttcatact	1860
ggtcaggctg	gtctcgaaact	cctgacccctc	agtgcacccac	ctggccttggc	ctcccaaagt	1920
gctgggattt	caggcgtag	ccaccacacc	cagccctcagt	tcccttattt	taataattgt	1980
cctatttgcgt	gggggttgcta	taaagattac	atgagacttt	aatggggaaag	tatttaggac	2040
agtgtatggg	acaaagttag	taccagacat	gcataaggca	tttggaaaacc	tgggtgcctt	2100
ctctctcact	cttaccattt	atagacaggg	gtcatataat	ttttttgtt	gagacggagt	2160
ctcgctctgt	ggcccaggct	ggagagcagt	gggtgtatct	cagctcatgg	gttcaagcga	2220
ttctctgtcc	tcagccctcc	gagtagctgg	gactataggc	atgcgccacc	aggcccagcc	2280
aatttttata	tttttagtag	agatgggggtt	tctccatgtt	agccaggctt	cccaggctgg	2340
tctcgatctc	ctgacccctcg	gtgatccacc	cacccggcc	tcccaaagt	ctgggattac	2400
aggcgtgagc	cactggccccc	tgccttctt	cacttttct	ttgagactgg	gtcgacacagt	2460
cgcacagtga	gactccatct	cctatggaaa	aaaagattgg	cctagcgtct	ggctcaggcc	2520
tgtatccca	gcactttggg	aggccttggc	ggccggatca	caaggtcagg	agatcgaggc	2580
cagcatggcc	aacatggtga	caccctgttt	ctactaaaaaa	tacaaaaaaa	attagccggg	2640
catggtggcg	catgcctata	gtgtcagctt	ctctggagaa	tcgcttgaac	ccgggaggccg	2700
gagggttgcag	tgagccgaga	ttgcgtcaact	gcactacagc	ctgggcaaca	gagcgagact	2760
cccgcttcag	aaaaaaaaaaag	ttttgggtgcc	aaccggacga	g		2801

<210> 39
<211> 2656

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 997947CB1

<220>
<221> unsure
<222> 2554, 2587, 2604, 2606, 2611, 2644, 2646-2647, 2651
<223> a, t, c, g, or other

<400> 39

```

ccactctcta gagctccccgg actcctagct ctgagaatgc ctgcggaaatg atcgcccccc 60
aggccggctg ccgcggctgc cgctgctgct gttattgcta ctgctgctgc cgccgcctct 120
gcttccactc ggctctgact ggcaggcaga aagtcaact gacgagggaa aggtctctgc 180
agttagtgg a ggcctacat aaaagagagt aaagaggggc aaaaacccag atcagaatgc 240
aggcgacgtc caaccttctc aacctcctgc tgctgtctt gtttgcggaa ttagatcctt 300
ccaagactca gattagtccct aaagaaggtt ggcagggtta cagctcagct caggatcctg 360
atggcgtgt cattgcaca gttgttgc tc cagaacaaaa cctgtgttcc cgggatgcca 420
aaagcaggca acttgcacaa ctactggaaa aggttcgaa catgtccca ag 480
tcttaaactt gagaactcag agagatccc aatatgttt aaaaatggaa acccaaatga 540
aagggctgaa ggcacaaatgg cggcaggatgg aagatgatcg aaagacactt atgaccaagc 600
attttcagga gttgaaagag aaaaatggacg agctcctgc tttgatcccc gtgctggAAC 660
agtacaaaac agatgctaag ttaatcaccc agttaacagga ggaataagg aatctgtctg 720
ctgtcctcac tggatttcag gaggaaatgg tgcctatgaa ctacggagaa ctacacaaaa 780
gagtgcttag cttgaaacaa agacttcgtg actgcatgaa aagactaaca tggcggaaac 840
tgatgaaaat cacaggccca gttacagtc a agacatctgg aacccgattt ggtgcttggaa 900
tgacagaccc tttacatct gggaaaaaca acagagtctg gtacatggac agttatacta 960
acaataaaaat tggtcgtaa tacaaatcaa ttgcagactt tgcgtggg gctgaatcaa 1020
ggacatacaa ctttccttcc aagtggggcag gaaactaacc tttgtctac aatggctcac 1080
tctattttaa caagtatcag agtaatatca tcatcaaata cagcttggat atggggagag 1140

```

tgcttgcaca acgaaggcctg gagtatgctg gtttcataa tgtttacccc tacacatggg 1200
 gtggattctc tgacatcgac ctaatggctg atgaaatcg gctgtggct gtgtatgcaa 1260
 ctaaccagaa tgcaggcaat attgtcatca gccaacttaa ccaagatacc ttggagggtga 1320
 tgaagagctg gagcaactggc taccggaaaga gaagtgcagg ggaatcttgc atgatctgtg 1380
 ggacactgtc tgcaccaac tcccaacttaa ctggagccaa ggtgttattat tcctattcca 1440
 cccaaacccat cacatgag tacacagaca ttcccttcca taaccaataat tttcacat 1500
 ccatgttgc ctacaatgca agagatcgag ctctctatgc ctggaacaat ggccaccagg 1560
 tgctgttcaa tgcacccctt ttccatcatca tcaagacaga ggtgacaca taggcaaatg 1620
 tgacatgtt tcattgattt aaacagtgtg atttgatataa gacccttcc 1680
 gttttttct tcactattat ttccatcat ttctccaaag caaagcattt ttattgtaaa 1740
 gttgggtgtt caaaaacata gctgagctt tctaacttac catgttggaa acacatctta 1800
 acttctaaat ttacaaggcc tattatgtcc ttgtcatgaa aagcactaaa aaaaaaaaaga 1860
 gtttaagtgg ctaaagtcat agttttgcaa gagattaatg atctgcctt tattagatc 1920
 agagactaat ggtgcttaa atgcacgaat gtctttttt tttaaaaactg tcattttta 1980
 ctgtcttttgc ctccatcatca gaaatattt tggttaggaat taggagaaca aaaagcactt 2040
 ttatcccatt tatttcttta aaaaatgtaa ggatttcatt tatattgaaa aataatatta 2100
 atcattttgc tgtaaacaca attctctgtat gcggtgctgt acagtcat 2160
 tgctaacatt ttatggcag tatgtatttc taccattgtt accaccattt tgctattgtt 2220
 tctcttcaat tctgtgaaag taatatttttataaaaatc actgaaattt aacctcagta 2280
 attgagtc ttttcaagtg tggtaagaa taatcttctt ggcttacccc ttacataag 2340
 cattataaaac taaaatgaaa accaaaccag acacctgaca tagatctttt attttacccc 2400
 aagtttttg ggccactgac attgaatgca acaactgatt tcatacaact gagttactct 2460
 gttcaacttca ctgaatgca cccatatagt ttcttgacca aggtgcatct ggattctaaa 2520
 tattggattt gagttgactc tactcattt ttanacacta aagaaatttt gttcttcata 2580
 gttaaangta cttagcattt attnanattt nacatacaac ttgcaataat gaaattcctt 2640
 atgncnngac nctgaa 2656

<210> 40
 <211> 968
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1521513CB1

<400> 40
 tctacatcta tcggagctga acttcctaaa agacaaaatg tttatcttca aagattcatt 60
 ctccctgaat cttaccaaca aaacactcct gaggagaaag aaagagaggg agggagagaa 120
 aaagagagag agagaaacaa aaaaccaaaag agagagaaaa aatgaattca tctaaatcat 180
 ctgaaacaca atgcacagag agaggatgtctcttccaaatgttctta tggactgttg 240
 ctggatccc catccattt ctcagtgcct gttcatcac cagatgtgtt gtgacattt 300
 gcatcttca aacctgtat gaaaaaaatg ttcagctacc tgagaatttca acagactct 360
 cctgctacaa ttatggatca gtttcagtca agaaatgttg tccattgaac tggaaatatt 420
 ttcaatccatc ctgtacttcc ttttctactg acaccatttccctggcggtta agttaaaga 480
 actgctcagc catgggggcac cacttggggtttatacaactt acaggaggag caggaatttcc 540
 tttctacaa gaaaactaaa atgagagatg ttttattgg actgtcagac caggttgc 600
 agggtcgtg gcaatgggtg gacggcacac cttgacaaa gtctctgagc ttctggatg 660
 tagggagcc caacaacata gctaccctgg aggactgtgc caccatgaga gactcttcaa 720
 acccaaggca aaatttggat gatgtaacct gtttctcaa ttatttcggttattgtaaa 780
 tggtaggaat aaatcccttg aacaaaggaa aatcttta agaacagaag gcacaactca 840
 aatgttaaa gaaggaagag caagaacatg gccacacccca ccccccacca cgagaaattt 900
 gtgcgctgaa ctccaaagga ctccataagt atttgttact ctgatataaa taaaataag 960
 tagttta 968

<210> 41
 <211> 1837
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1863994CB1

<400> 41	ccccagcccc	gctcagtccc	gagcgccccg	agtgcgtcg	ccgcccgcgc	aagcatgcag	60
taaaggctga	aatctgggt	cacagctgag	gaagaccta	gacatggagt	ccaggatgt	120	
gcctgcgtg	ctgtgtccc	acctcctccc	tctctggca	ctgctgttgc	tgccccctccc	180	
accgcctgt	cagggtctt	catcctcccc	tcgaacccc	ccagccccag	cccgcccccc	240	
gtgtgccagg	ggaggccccct	cggcccccacg	tcatgtgtgc	gtgtgggagc	gagcacctcc	300	
accaagccga	tctcctcggg	tcccaagatc	acgtcggcaa	gtccctgcctg	gcactgcacc	360	
cccagccacc	ccatcaggct	ttgaggaggg	gccccccta	tcccaatacc	cctgggctat	420	
cgtgtgggt	ccccaccgtgt	ctcgagagga	tggagggac	cccaactctg	ccaatccccgg	480	
atttctggac	tatgtttttt	cagcccccta	tgggctcgca	acccccacacc	ccaaactcaga	540	
ctccatgcga	ggtgatggag	atgggcttat	ccttggagag	gcacctgcga	ccctgcggcc	600	
atccctgttc	ggggccctgt	gggaagggtgt	ggaccccccag	ctctatgtca	caatttacat	660	
ctccatcatc	atttgttctcg	tggccactgg	catcatcttc	aagttctgt	gggaccgcag	720	
ccagaagcga	cgcagaccct	cagggcagca	aggtgcctg	aggcaggagg	agagccagca	780	
gccactgaca	gacctgtccc	cggctggagt	cactgtgtc	ggggccttc	gggactcacc	840	
taccccccacc	cctgaccatgt	aggagccccc	agggggaccc	cggcctggga	tgccccaccc	900	
caagggggct	ccagccttcc	agttgaaccg	gtgagggcag	gggcaatggg	atgggagggc	960	
aaagagggaa	ggcaacttag	gtcttcagag	ctgggggtgg	ggtgcctct	ggatgggtag	1020	
tgaggaggca	ggcgtggcct	cccacagccc	ctggccctcc	caagggggct	ggaccagctc	1080	
ctctctggga	ggcaccccttc	cttctccca	tcttcagga	tctgtgttct	attctctgtct	1140	
gcccataact	ccaaactctgc	cctcttggt	tttttctcat	gccaccttgt	ctaagacaac	1200	
tctgcctct	taaccttgtat	tcccccttct	tgtcttgaac	ttcccccttct	attctggcct	1260	
acccttggt	tcctgactgt	gcccccttcc	tcttccttc	aggattcccc	tggtgaatct	1320	
gtgatggcc	caatgttggg	gtcagccaa	gcaggaggcc	aagggggccgg	cacagcccc	1380	
atcccactga	gggtggggca	gctgtggga	gctggggca	caggggctcc	tggctctgc	1440	
cccttgacca	ccaccccgaa	cactccccag	ccccacgggc	aatccatat	gtcgccccct	1500	
ctgcagggtt	gggcctcaca	tatctgtgac	ttcgggtccc	tgtccccacc	tttgtgcact	1560	
cacatgaaag	ccttgacac	tacacccac	ttcacacaggc	catttgacaca	cgctcctgca	1620	
ccctctcccc	gtccataccg	ctccgctcag	ctgactctca	tgttctctcg	tctcacat	1680	
gcactctctc	cttcccacat	tctgtgtc	gtcactcag	tggtcagcgt	ttccctgcaca	1740	
cttacccct	catgtgcgtt	tcccgccctg	atgttgcgtt	ggtgtgcggc	gtgcactcactc	1800	
tctccctcat	gaacacccac	ccacccctcg	tcgcagc			1837	

<210> 42
<211> 2124
<212> DNA
<213> *Homo sapiens*

<220>
<221> misc_feature
<223> Incyte ID No: 2071941CB1

```

<400> 42
gtgacgcggc tgccggaggtg acgcgggagg tcgcgcgccc cttccggcgc ggggaggggcg 60
ctgaagatcg gggccgctcg gccgcaggcg cctccagcgc cgccggatgt agcgcggggg 120
accgcggccc ccagcagagc ccgcctgccc ggcttgccta ccatcagagg gagatctctg 180
ccccctgggg ctgagagacc ccaacccccc cccaaagctga agctgcaggg tattgaggt 240
ccagccagat gtcttccac aaaggatctg tggtggcaca ggggaatggg gctccgtcca 300
gtaacagggg agctgacacg gtggaaactgg ctgaactggg accccctgccta gaagagaagg 360
gcaaaccgggt aatcgccaaac ccacccaaag ctgaagaaga gcaaaccatgc ccagtgc 420
aggaagaaga ggaggagggtg cgggtactga cactccccct gcaagccac cacggcatgg 480
agaagatgga agagtttgtg tacaaggctt gggagggacg ttggagggtc atcccatatg 540
atgtgtccct tgactggcta aaggacaacg actatctgtc acatggctt agacccatcca 600
tgcccttcct tcgggcttgc ttcaagagca tcttccgcatt ctcatacagaa actggcaaca 660
tctggaccca tctgttggt ttctgtctgt ttctttttt gggaaatctt accatgtct 720
gaccaaatat gtacttcatg gcccccttac agggagaaggt gttttttggg atgttctttt 780
tgggtgcagt gctctgcctc agcttcttct ggcttttca caccgtctat tgtcattcag 840
agaaaagtctc tcggactttt tccaaactgg actattcagg gattgtctt ctaattatgg 900
ggagcttgc cccctggctc tattattct tctactgct cccacagcca cggctcatct 960
acctctccat cgctctgtgtc ctgggcattt ctgcctatcat tttttttttt tttttttttt 1020
ttgccactcc taagcaccgg cagacaagag caggcgtttt cctgggactt ggcttgatgt 1080
gcgtctgtcc caccatgcac ttactatcg ctgagggtt tttttttttt tttttttttt 1140
gccagatggg ctggttcttc ctcatggctg tgatgtacat cactggagct ggcctttatg 1200

```

ctgctcgaat tcctgagcgc tccttcctg gaaaatttga catatggttc cagtctcatc 1260
 agatttcca tgtcctggc gtggcagcag ccttgcataa gtcctccaacc 1320
 ttcaggaatt ccgttacggc ctagaaggcg gctgtactga tgacaccctt ctctgagcct 1380
 tcccacctgc ggggtggagg aggaacttcc caagtgcctt taaaaataaac ttcttgcgt 1440
 aagtggagg aagactctga gttgtctgt tctagaagaa acctcttgc gaattcgt 1500
 ccaaccaagg ttcagccac ttgcacaccc actggcaat aaacttcca ttccattct 1560
 cctagctgg gatggggcat ggtcaaaactt agccatcccc tcctcagcaa ggcatttacc 1620
 ggcgcctcac agagacgta ctttgcactt catgttgcata ttttaccctt tcctccaacc 1680
 attttggaa aattatggac tggacttcc cagaaattct gtctttctt ctggaaagaaa 1740
 atgtccctcc cttacccccc tccttaactt tgcatacttgc ttataacag gccatccatt 1800
 ttgttagcac actttcaaa aacaattata taccctggc ccatcttctt agggcctgga 1860
 tctgttata gaggcaggaag aataaaagcca ccaactttt cctagcccg ctaatcatgg 1920
 aagtgtgtcc aggcttcaag taacttgcgt ttaattttt tttttttct tggcagagta 1980
 atgtaaaatt taaatggggaa aagatattta atattaata ctaagcttta aaaagaaaacc 2040
 tgctatcatt gctatgtatc ttgtgcataa gactatgttgc ttaataaaaag aaagtacaga 2100
 agagacttgg cattcaaaaa aaaa 2124

<210> 43
 <211> 993
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2172512CB1

<400> 43
 ccgcgcgtt actttgttta taacttgcataa aatcctctcc gtctcccttc cctgcctcct 60
 ttcccttccc ttccctctgc cagtacaact agaccggcg tctggcgccc cccgtggccca 120
 gcattctgcg gggcaggcg attaatttgc attttcaaaatgtcagggt tggtacccac 180
 agccctgaa cggcctgcg gtggaaatggaa aatccaaaca aaaccaccag atccaaaggcc 240
 tgcattcttcc cttgttataca gttcttgcattt ttaccaggaa cccctggaa cagctgtccc 300
 tccacccatc ggcgttcccg gaggcttgc tatgggatc tacagtccac agcaacccag 360
 taccccttcc ttgttaccgc cagttgggg tatccatctt gtccggatc agcctggccaa 420
 atatccatgc ccaaatcgtt ctgttccat aacatggatc ccaggccaa ctcctatggc 480
 aaactccctt cctggcttgc aataacttagt tcagttggatc aacatacatg ttcttcagca 540
 ttttggccctt ctggaaatgtt tgacatgttt tgaaactaat aatagatatg atattaaaaaa 600
 caactcagac cagatggttt acattgttac cgaagacaca gatgacttta ccaggaatgc 660
 ctatccgaca ctaaggccct tcgttcccg ggttactgtat tttatggggc gagaatcat 720
 gacaatgcg agacccttca gatgcacccgt ctgttgcctt ttgttgcctt ctgcccagaca 780
 agagctggag gtgcagttgc ctccctgggtt caccattggc ttgttgcgg aacattggaa 840
 cctgtgcagg gcgggtgtaca gcatccaaaa aaaaaaaaaaaa aaaattggcgcc cgcaagctta 900
 ttccctttagt tgaggttaat ttttgggggc atgtgcctgt gcagtgcggg ggttgggtt 960
 gttgtgtggg ggtgggggtgg ggggttgggtt ggt 993

<210> 44
 <211> 2214
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2483172CB1

<400> 44
 ggctctcgcc ccagcgcgcgccc tgccttcgccc gcccggcgcc gtcctcgcc cgctgcacga 60
 cgacgcgcacg cccctgtgc aggccggcga cccggaccggc cccagaccca gacgcacccat 120
 ggcgcacggcc gctgtactgc ctcagctgc cccagctggc ctccggatgc acctacggac 180
 tgactgtggg ggcagagaag ggcgagatca ggactctgtc ttgtttaatc gtgactgcatt 240
 gaaggtcgcc tccctcgccc ctacttgggtt ggagtgtctt gtattgttct aaggccagga 300
 gcacgggttag ccacagtctg ttggtagaat ttggcgttctt gatagttgg aaaaatggcgaa 360
 tgacactgtt ggaagacttgg tggatgttgc ctcccagaga gctctgttag 420
 tctggggcat cccagtgaaac tgcgttggagg ctggaaatcga agagaccctc caggctgcga 480

tgcccccaagg	ctccattaccga	atgcttggga	gaatgttctg	gagggaagaaa	aatgcgaaag	540
cagccttatt	agagctca	ggcgctgtag	attacccgc	gatccccagg	gagatccgg	600
gcaaaggagg	ggtctggaaa	gtgttattta	agcccccaac	ttctgtatgt	gaatttttag	660
aaagattgca	cctcttccta	gctagagagg	ggtgaccgt	gcaagatgtt	gcccgtgtcc	720
ttgggttca	gaaccctact	ccgaccccccgg	gccagagat	gccagcagag	atgctaaact	780
atattttgg	taatgttatt	cagcctcttg	ttgagtccat	atggtacaag	aggctgacac	840
tttctcggg	gaggacatc	ccagggcctg	gagagggaaac	cttgcattccc	tggctggagc	900
acactaatga	ggtcttagag	gagtggcagg	tgtcgtatgt	agaaaaagagg	cgcggttga	960
tggagagtt	tagaggcccc	gccgctgtatg	ttatgcat	cctaagtcc	aacaaccccg	1020
cgataaccac	tgcgaatgc	ctgaaggcgc	ttgacgggt	gttggggagc	gttggagact	1080
ctagggatgc	ccagatcaa	tttctgaaca	cttacgaaa	cccgggagaaa	aaattgtctg	1140
cttatgtcat	tgcgtctgg	cctctgtctg	agaaggtgg	agagaagggg	gccattgata	1200
aagataatgt	gaaccaggcc	cgcttagagc	aggtcattgc	cgggggccaa	cacagcgggg	1260
ccatccgaag	gcagctgtgg	cttacggggg	ctggggagg	gccagccca	aaccttttc	1320
agttgtctgt	gcagatccgt	gaggagggaa	ccaaggagga	ggaggaggag	gctgaggcca	1380
cccttctgc	gttaggcctg	gaagggcact	tctgatgtcc	aggaaaggca	gttttagtgc	1440
agaccttagat	cacagctact	tttcttgc	ctgtgggtc	ttacagatgt	gtctctgagt	1500
agtaaaggct	tagccttgg	ctgttttgg	gtttttgg	ggggaaagg	agtcaggcct	1560
gagttattcat	gtaacattct	aaaattgtgc	cagcgagcac	cgtgaacgac	tgcaatgca	1620
gcgggtcttg	ctggctaaagg	taaagggtt	gttgacaca	gctgttagt	cacgtgtca	1680
tcatggacat	cataatcagt	tgtaaaaac	acgcgaacct	atgacacttc	ttattccaca	1740
ctgaatgtga	aattgcattgt	tcatgtt	actacgaggc	ctggctcaca	ggaagtgtt	1800
agtaaaagta	tgcactgtta	gattactgt	aacgcggata	gattttttgtt	taccataat	1860
tgttccagat	ttatattat	ggaaggaaat	gtgcatttt	tagtattac	tcaactttac	1920
aatgcaaaaa	tcttattttct	catctttaaa	catgtcgacc	agtttaattt	aaaagtattc	1980
tgagactgca	aaatggggtg	ttaaaaaaa	ctgcagttac	ggagctgtgt	aaaccagttt	2040
ctcatttgc	aaatgcata	tgtaaattgc	atggagaggt	tgatatgcac	ctgtacagta	2100
attactcc	ccatttcaca	tcttgcag	agaatagttc	ttgttcatac	tgagtgttct	2160
aaatttgaag	ttatataatac	aaattaaaaat	atttaaaaa	ttcaaaaaaa	aaaa	2214

<210> 45
<211> 897
<212> DNA
<213> *Homo sapiens*

<220>
<221> misc_feature
<223> Incyte ID No: 2656128CB1

```

<400> 45
gcaaaaatgca tgacagtaac aatgtggaga aagacattac accatctgaa ttgcctgcaa 60
acccagggtt tctgcattca aaagagcatt ctattaaagc taccttaatt tggcgcttat 120
ttttcttaat catgtttctg acaatcatag tgggtggaaat gttgtctgt ttaagcgcaa 180
taagagctaa ctggccatcaa gagccatcag tatgtctca agctgcatgc ccagaaagct 240
ggattgggtt tcaaagaaaag tgggttctatt ttctgtatga caccagaagac tggacatcaa 300
gtcagagggtt ttgtgactca caagatgctg atcttgcctca gttgtggaaagc ttccaggaac 360
tgaatttcc tggagatataa aaggcccat ctgatctact gatggggctg agcagagaaac 420
aaggcccaacc atgaaaatgg ataaatggta ctgaatggac aagacatgtt gtcatgaaag 480
aagatgtgc caactgttat gttgcaaaagg ttccacaagt tcctcgaatg aatccaaagac 540
ctgtcatgtt ttccatctc gggagcagga gagttgtgcct atttgaatgta caaagggtgcc 600
agtagtgcca ggcactacac agagaggaaag tggattttgtt ccaaatcaga tatacatgtc 660
tagatgttac agcaaaagccc caactaatct tttagaagcat attggaaactg ataactccat 720
ttaaaaatga gcaaaaattt tattttttt accaacaggt atatgaaaat atgctcaata 780
tcactaataa ctggggaaaat acaaaatcaaa atcatagtaa aatattacact gttttcatgg 840
tgctaatatt acctgttctc ccactgtctaa tgacatacccc gagactgagt aatttaa 897

```

<210> 46
<211> 2167
<212> DNA
<213> *Homo sapiens*

<220>
<221> misc_feature

<223> Incyte ID No: 5855841CB1

<400> 46

<210> 47

<211> 1235

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 603462CB1

<400> 47

gtggagccgg	cgagagagtgc	gcagccccgg	ctgatggaa	tcgcgtgggg	gtcgagagg	60
gcaccctact	gtatccagca	tgctccaagg	ccacagctc	gtgttccagg	ccttgcgtgg	120
gaccccttc	acctggggga	tgacagcagc	tggggcagct	ctcgtgttc	tattctctag	180
tggacagagg	cggatcttag	atggaaagtct	tggcttgct	gcaggggtca	tgttggcagc	240
tccctatgg	tctcttctgg	ccccagca	tgagatggcc	acgttcctcg	ggggcttcgg	300
tgccttgcc	ttctccctcg	tggctgttgg	cttcaccccg	ggagccggct	ttgttctactt	360
ggctgaccc	ctgatgcctc	acttgggtgc	agcagaagac	ccccagacgg	ccctggcact	420
gaacttcggc	tctacgttga	tgaagaagaa	gtctgtatcct	gagggtcccc	cgctgtctt	480
ccctgagagt	gaactttcca	tccggataga	caagagttag	aatggtgagg	catatccagag	540
aaaagggcg	gcagccactg	gggttccaga	gtccctgtgc	cttctcgagg	600	
gaatctggca	cagccccggcg	gcagcagctg	gaggaggatc	gcactgtcta	tcttggccat	660
cactatacac	aaagtccag	agggtctcgc	tgttggagtt	ggatttgggg	ctataaaaa	720

gacggcatct gctaccttg agagtccag gaattggcc attggaatcg ggatccagaa 780
 tttcccgag ggcctggctg tcagcctcc cttgcgaggg gcaggcttct ccacctggag 840
 agcttctgg tatggcagc tgagccgc ggtggagccc ctggccggg tctttgggtgc 900
 ctggcctg gtgctggctg agcccatctt gcccacgc ctggcccttgc ctggccgtgc 960
 catggtctac gtggcatgg acgacatcat cccgaagcc cagatcagtg gtaatggaa 1020
 actggcatcc tggccctca tcctggatt tggatgtat atgtactgg acgttggcct 1080
 gggctaggc tgagacgctt cggacccgg gaaaggccat acgaagaaac agcagtggtt 1140
 ggctctatg ggacaacaag ctttttctt cacattaaaa ctttttctt tcctcttc 1200
 ttcatctcat ttccttgatt gactctaatt ataat 1235

<210> 48
 <211> 2257
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 747681CB1

<400> 48

ccgagccgag cggaccgaag gcgcgcccga gatgcaggta agcaagagga tgctggccgg 60
 gggcgtgagg agcatgccc gccccctctt ggcctgtgg cagcccatcc tcctgttgt 120
 gctggctca gtgctgttag gtcggccac gggctgccc ccccgctgcg agtgcctgc 180
 ccaggaccgc gctgtgtgt gccaccgcaa gcgccttgcg gcagtcctcc agggcatccc 240
 caccgagacg cgcctgtgg acctaggcaa gaaccgcata aaaacgctca accaggacga 300
 gttcgcacgc ttcccgacc tggaggagct ggagctcaac gagaacatcg tgagcgcgt 360
 ggagcccgcc gccttcaaca acctttcaaa cttccggacg ctgggtctcc gcagcaaccg 420
 cctgaagctc atccccgttag ggcgtttcac tggcctcagc aacctgacca agctggacat 480
 cagcgagaac aagattgtta tcctgtggta ctacatgttt caggaccgt acaacctcaa 540
 gtcacttggag gtggcgaca atgacatcgat ctacatctt caccgcgcct tcagcggcct 600
 caacagcctg gaggcagctg cgctggagaa atgcaacatcg acctccatcc ccaccgaggc 660
 gctgtccac ctgcacggcc tcatcgctt gagctccgg cacattcaaca tcaatgccat 720
 cgggactac tccttcaaga ggcttacccg actaaaggctt ttggagatct cccactggcc 780
 ctacttggac accatgacac ccaactgcct ctacggcctc aacctgacgt ccctgtccat 840
 cacacactgc aatctgaccg ctgtccctta cttggccgtc cggccacccat tctatctccg 900
 ctccctcaac ctcttcttaca accccatccg caccatttagg ggctccatgt tgcatgagct 960
 gctccggctg caggagatcc agctgggtgg cggcagctg gccgtgggg agccctatgc 1020
 ctcccgccgc ctcaactacc tgccgtgtcaatgtctt caccgcgcct ggcaaccgc tgaccacact 1080
 ggaggaatca gcttccact cgggtggccaa cttggagaca ctcatcttgc actccaaccc 1140
 gctggcctgc gactgtcgcc tcctgtgggt gttccggcgc cgctggccgc tcaacttcaa 1200
 cggcagcagcc cccacgtgcg ccacgcggaa gttgtccag ggcaaggagt tcaaggactt 1260
 ccctgatgtc tcactgcccactacttcac ctggccgcgc gcccgcattcc gggaccgc 1320
 ggcggcagcg gtgtttgtgg acgaggccca caccgtgcag tttgtgtgc gggccgtatgg 1380
 cgaccggccgc cccgcctcc tctggcttc acccccgaaat caccgtgtc cagccaagag 1440
 caatggccgc tcacatcgatcc tcctgtatgg caccgtggag gtgcgtacg cccaggata 1500
 ggacaacggc acgtacatgt gcatcgccgc caacgcggggc ggcaacgact ccatgcccgc 1560
 ccacctgcat gtgcgcagct actcgccgaa ctggcccccatt cagcccaaca agacccatgc 1620
 tttcatctcc aaccagccgg gcgaggggaga ggccaacagc acccgccca ctgtgcctt 1680
 cccctcgac atcaagaccc tcacatcgcc caccaccatg ggcttcatct ctccctgggg 1740
 cgtcgccctc ttctgcctgg tgctgtgtt tctctggagc cggggcaagg gcaacacaaaa 1800
 gcacaaacatc gagatcgagt atgtccccg aaagtccggac gcaggcatca gctccggcga 1860
 cgcgcggccgc aagttcaaca tgaagatgtat atgaggccgg ggcggggggc agggaccccc 1920
 gggcggccgg gcaggggaag gggcctggcc gcoacctgtc cactctccag tccttccac 1980
 ctcccttcttca cccttcttaca cacgttctt tcctccctcc cgcctccgtc ccctgtgtcc 2040
 ccccgccagc cctcaccacc tgccctccctt ctaccaggac ctcagaagcc cagacctggg 2100
 gacccacactt acacaggggc attgacagac tggagttgaa agccgacgaa ccgacacgcg 2160
 gcagagtcaa taattcaata aaaaagttac gaacttctc tgtaacttgg gtttcaataa 2220
 ttatggattt ttatggaaaac ttgaaataaaa aaaaaaaaaa 2257

<210> 49
 <211> 2359
 <212> DNA
 <213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 919469CB1

<400> 49

gtgggatttg aagatccact ccacttctgc tcatggcggtt ccagggcctg cccctgcacg 60
 tggccacact gctgactggg ctgctggaaat gcctgggctt tgctggcgctc ctcttggct 120
 ggccttcaact agtgttgc ttcaagaatg aagattactt taaggatctg tgtggaccag 180
 atgctgggcc gattggcaat gcccacaggc aggctgactg caaagcccaag gatgagaggt 240
 tctcaactcat cttcacccctg gggtccttca tgaacaactt catgacattt cccactggct 300
 acatcttga ccgggtcaag accaccgtgg cacgcctcat agccatattt ttctacacca 360
 cccgcacact catcatagcc ttcacctctg caggctcagc cgtgtcgctc ttctggcca 420
 tgccaatgtt caccattggg ggaatcctgt ttctcatcac caacctgcag attggaaacc 480
 tatttggcca acacccgtcg accatcatca ctctgtacaa tggagcattt gactcttctt 540
 cggcagttt ccttatttt aagcttctt atgaaaaagg catcagcctc agggcctcct 600
 tcatacttcat ctctgtctgc agtacactggc atgtacgtacg cactttcctc ctgtatcccc 660
 gggggcactt cccataccctt ctgcggggca actacagctt tggcctgtgc cctggaaatg 720
 gcaccacaaa ggaagagaag gaaacagctg agcataaaa caggagacta cagtc当地agg 780
 agttccttc agcgaaggaa gagaccccaag gggcaggcga gaagcaggaa ctccgctcct 840
 tctggagcta cgcttctctt cggcgttgc cctggcacct ggtgtggctg tctgtgatac 900
 agttgtggca ctaccttctt attggcactt tcaactcctt gctgaccaac atggccgggt 960
 gggacatggc acgagtctgc acctacacaa atgccttgc cttcaacttgc ttcggagtgc 1020
 tttgtgcccc ctggaaatggc ctgctcatgg accggcttac acagaagtac cagaaggaag 1080
 caagaaagac aggttccctcc accttggggg tggccctctg ctgcacgggt cttcgctgg 1140
 ccctgacatc cctgtgtgc ctgggcttcg ccctctgtgc ctcaatcccc atccctccctc 1200
 tccagttactt cacccatcatc ctgcaagtga tcagccgtc cttccctctat gggagcaacg 1260
 cggccttctt cacccttgc ttccctttagt agcactttgg caagctctt gggctggta 1320
 tggccttgc ggtgtgggt tctctgttcc agttcccatt cttcacccctc atcaaaggct 1380
 cccttcagaa tgaccattt tacgtgaatg tgatgttcat gcttgccattt ctctgtacat 1440
 tcttcaccc ctttctggta tttcgaaaatg gccgtacttg gaaagaaaatg ccctctgcaa 1500
 ttgcataatgtt cagaagccctt cactttttagt ccccgaggat ggtttttgttcc atcttccacc 1560
 accttggagg acctgtgtgc cccaaagact ttgcctatcc cagcaaaaaca cacacacaca 1620
 cacacacaca cacaaaataa agacacacaa ggacgtctgc gcagcaagaa aagaatctca 1680
 gttgccaagc agattgatatac acacagact caaagcaag gcatgtggaa cttctttattt 1740
 tcaaaaacaga agtgccttgc ttgcacttagc ctggcagac cttgactcc agggagatg 1800
 accttggggg ggaagtgtgtt caactatttc tttaggctgtt tttggctccg aagccttatat 1860
 gtgcctggat cctctgccc gggttaaattt ttcaaggtaa gagtgaggtt gtcatggct 1920
 cagctatgtt tcctggctct ccctcaagag tgcaaggctt gctagagaac tcacagctct 1980
 gggaaaaaaa ggacgcacaca gggttccctg ggcccaggctt cagcccaagcc actgtatgt 2040
 gatgaccccttgc gcctgaccctt ggtctggctt cagaatactt ttccctcatctt gtaaaattga 2100
 gatgaattttt ggtgtggaaa ttcccttgc gggcggatgtt cttttttttttt gtttaggaata 2160
 gtgacagatg caggccaccc caagggccat gggagccagc tgacctgtctt gaccgaagga 2220
 tttctgacag actatctttt gggatgtttt caagaaggaa tataagttat ttactttggg 2280
 cattttaaatggaaa aaaaatttctc tcgggaaataaa ttttatagaa aaataaaagct tctgtgtctt 2340
 aggcaaaaaaa aaaaaaaaaa 2359

<210> 50

<211> 2052

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 977658CB1

<400> 50

gggtttgaag ccgcgcgcgc agggagcgag gtcgcagtga cagcggcggtt cgatcgacc 60
 caggctggcc cggcgatccc gcctgcgtcc cgcgcctcccg ccccaagcatg acagccccgg 120
 cgggtccgcg cggctcagag accgagcggtt ttctgacccca caaccccggtt tatgggaccc 180
 aggcggggcc ttccacccggcc cttccgcacac ccccaagaaga ggaagacctt cggcgctgtc 240
 tcaaataactt ttcatgagt ccctgcgaca agtttcgagc caagggccgc aagccctgca 300
 agctgtatgtt gcaagtggc aagatccctgg tggtcacgggt gcaagctcattt ctgtttggc 360
 tcagtaatca gctggctgtt acattccggg aagagaacac catgccttc cgacacctct 420

tcctgctggg ctactcgac ggagcggatg acaccccgac agcctacacg cgggagcgc 480
 tgtaccaggc catcttccat gctgtggacc agtacccgtgc gttgcctgac gtgtcaactgg 540
 gccggatgc gtatgtccgt ggtgggggtg acccttggac caatggctca gggcttgctc 600
 tctggccacg gtactaccac cgaggccacg tggaccggc caacgacaca tttgacattg 660
 atccatgtgtt ggtactgac tgcattccagg tggatccccc cgagcggccc cctccgcccc 720
 ccagcgcacga tctcacccctc ttggaaagca gtcaggatca aagaacccatc acgctcaat 780
 tccacaagct ggtcaatgtc accatccatc ctcggctgaa gaccattaac ctccagagcc 840
 tcataataa tgagatcccg gactgctata ccttcagcgt ctcgtacacg tttgacaaca 900
 aagcacacag tggccggatc cccatcagcc tggagaccca gcccacatc caggagtgt 960
 agcaccacccat ggtcttccag cacggagaca acagcttccg gtcctgttt gacgtgggt 1020
 tcataccctac ctgctccctg tccttcctcc tctgcggccg ctcacttcctt cgaggcttcc 1080
 tgctgcagaa tgagtttggg gggttcatgt ggcggcagcg gggacgggtc atcaggctgt 1140
 gggagcggct ggaatttgc aatggctgtt acatccctgtc cgtcaccaggc gatgtgtca 1200
 ccatctcggg caccatcatg aagatccggca tcgaggccaa gaacttggcg agtacgacg 1260
 tctgcagcat ctcctcgggc acctcgacgc tgctgggtgt ggtggcgtg atccgttacc 1320
 tgaccccttcc cccaaactac aatatccctca tcgcccacact gcccgggtggcc ctgcccacgc 1380
 tcatacgctt ctgctgctgc gtggctgtca tctacccctggg ctactgcttc tgggtgtga 1440
 tctgtgtggg gccctatcat gtgaagttcc gtcacttc catgggtgtct gagggtgtgt 1500
 tctcgctcat caatggggac gacatgttt tgacgttgc cggccatgcag ggcacggcagg 1560
 gcccacggcag cctgggtgtgg ctcttcctcc agcttaccc ttaacttcctt atcaggcttc 1620
 tcatactacat ggtgtcagc ctcttcatcg cgctcatcac cggcccttac gacaccatca 1680
 agcatcccg cggccacggc gcagaggaga gcgagctgca gcccatacatc gcacagtgc 1740
 aggacagccc caccccccgc aagttccggc gcggggagcgg ctcggcctgc agccttctt 1800
 gctgtgtggg aaggggacccc tcggaggagc attcgctgt ggtgaattgt ttcgacactga 1860
 ctggcgttgg accgttaggcc ctggactgca gagacccccc ccccccggcc cgttattttt 1920
 tttgttaggtt ttgttttaa ggatccggctc cctgtcgcc cccggggggg cctggacctt 1980
 tcgtgtcgga cccttggggg cggggagact ggggggggag ggtgttgaat aaaaggaaaa 2040
 ataaaaaaaaaa aa 2052

<210> 51

<211> 1939

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1004703CB1

<400> 51

cttgtatggcg tcgggctgga gagccgcagt cccggctgca gcacccggaa gaaggcagac 60
 cgtgtgggg ggcctgtggc ccagcgtgt gtggcctcg ggagtggaa gtggaggcag 120
 gagcccttcc tacacttcgc catgagttt ctcatcgact ccagcatcat gattacctcc 180
 cagataactat tttttggatt tgggtggctt ttcttcatgc gccaatttgg taaagactat 240
 gagatacgtc agtatgtgtt acaggtgtc ttctccgtga cgtttgcatt ttcttgccacc 300
 atgtttgagc tcatacatctt tgaatctta ggagtttgc atagcagctc ccgttattttt 360
 cactggaaaa tgaacctgtg tgaatttgc ctgatccctgg ttcttcatggt gccttttac 420
 attggctatt ttattgttag caatatccga ctactgcata aacaacgact gctttttcc 480
 tgcctcttat ggctgacccct tatgtattt tcctggaaac taggagatct ctttcccatt 540
 ctcagcccaa aacatggggat ctatccata gaacagctca tcagccgggt tgggtgtgatt 600
 ggagtgaact tcatacgctt cttttcgttgg tttggtgcgt tcaactgcctt atacacttac 660
 atgtcttact tcctcaggaa tgcgtactgac acggatattc tagccctggg acggcactg 720
 ctgcaaccca tggatatgt cataagcaa aagaaaagga tggcaatggc acggagaaca 780
 atgttccaga agggggaaat gcataacaaa ccatcagggtt tctggggaaat gataaaaaagt 840
 gttaccactt cagcatcagg aagtggaaat ttactctta ttcaacagga agtggatgt 900
 ttggaaat taagcaggca gcttttcgttggaaacagctg atcttatatgc taccaggag 960
 agaatagaat actccaaaac ctcaagggg aaatatttttta attttcttgg ttacttttc 1020
 tctattttact gtgtttggaa aattttcatg gtcacccatca atattttttt tgatcgatgt 1080
 gggaaaacgg atccctgtcactt aaggccatc gagatcactg tgaattatgtt gggaaatccaa 1140
 tttgtatgtt gttttttggc ccaacacatt tccttcattt ttgttggaaat aatcatcgcc 1200
 acatccatca gaggattgtt gatcactt accaacttctt tttatccat ctcttagcagt 1260
 aagtccctca atgtcattgtt cctgtatataa gcacagataa tggccatgtt ctttgcctcc 1320
 tctgtgtgc tgcattccat gatgtatgtt ttagaataacc gcaccataat cactgaagtc 1380
 cttggagaac tgcagttcaa ctcttatcac cgttggttt atgtgtatctt cctggcagc 1440

gctctctcta gcatactctt octctattt gctcacaaac aggcaccaga gaagcaaatg 1500
 gcaccttggaa ctttaaggcta ctacagactg ttagaggcca gtggttcaa aatttagata 1560
 taagaggggg gaaaaatgga accagggcct gacatttat aaacaaacaa aatgctatgg 1620
 tagcatttt caccctcata gcatactct tccccgttag gtgatactat gaccatgg 1680
 agcatcagcc agaacatgg agggagaact aactcaagac aataactcagc agagagcata 1740
 ccgtgtggat atgaggctgg ttagaggcg gagaggagcc aagaaaactaa aggtgaaaaa 1800
 tacactggaa ctctggggca agacatgtt atggtagctg agccaaacac gtaggatttc 1860
 cgtttaagg ttcacatgga aaaggttata gcttgcctt gagattgact cattaaaatc 1920
 agagactgta aaaaaaaaaa 1939

<210> 52
 <211> 1138
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1334051CB1

<400> 52
 caaactgaa cttatatctg caatttatgg tggtagatgc aagaggtatg ccagtagcac 60
 actggtgct tcagaagaaa ttctcaacac ctagctcgcc agagagtcta tgtatggat 120
 tgaacaatct gtaaactaaa ggatcctaattt catgaaaata agtatgataa attataagtc 180
 actattggca ctgttggat tattagccctc ctggatcatt tttacagttt tccagaactc 240
 cacaaaggtt tggctctgctc taaacttatac catctccctc cattactggaa acaactccac 300
 aaagtcccta ttccctaaaaa caccactgtt atcattaaag ccactaacaag agactgaact 360
 cagaataaaag gaaatcatag agaaaactaga tcagcagatc ccacccagac ctttcaccca 420
 cgtgaacacc accaccagcg ccacacatag cacagccacc atcctcaacc ctcgagatac 480
 gtactgcagg ggagaccagc tgacatcct gctggagggtt agggaccact tgggacgcag 540
 gaagcaataat ggcggggatt tcctgagggc caggatgtct tccccagcgc tgatggcagg 600
 tgcttcagga aagggtactg acttcaacaa cggcacctac ctggtcagct tcactctgtt 660
 ctgggagggc caggctctc tgcactccat cccagtgaag ggggtgcage 720
 tctctggagt gcaaggaacc aaggctatga cagggtgatc ttcaactggc agtttgcataa 780
 tggcaactcc caagtcact ctgaatgtgg cctgatccata aacacaatg ctgaattgtg 840
 ccagtcactg gacaacagag accaagaagg cttctactgt gtgaggcctc aacacatgcc 900
 ctgtgctgca ctcactcaca tgcattctaa gaacaagaaa gtttcttatac ttagcaaaca 960
 agaaaagagc ctcttgaaa ggtaaaaata attacttctt gagactacatc gtgcaaatat 1020
 tgcattttgg cctatataact gatccaaaga aaagtcttgc gatgttatta atttgggtg 1080
 tcttttagaa gagccttgg ggaaaggatc tgtgaattca ttttagagaca gtgccccat 1138

<210> 53
 <211> 2117
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1336728CB1

<400> 53
 tggccacacgc aacacccact gaggcacgtg ggagctgagt atggcgtccc tggctcgct 60
 ggagctgggg ctgtttctgg ctgtgctgtt ggtgacggcg acggcgtccc cgcctgctgg 120
 tctgctgagc ctgcacccatc ctggccaggcg cgctctggat caagaggctc tggcggcct 180
 gttaaatacg ctggcggacc gtgtcactg caccaacggg ccgtgtggaa agtgcctgtc 240
 tgcggaggac gcccctggcc tggcgagcc tgagggtca gggctgcccc cgggccccgt 300
 cctggaggcc aggtacgtcg cccgcctcag tgccgcggcc tccctgtacc tcagcaaccc 360
 cgagggcacc tgcggggatc ctcggctgg cctctggcc tccatgcag accacccct 420
 ggcctgtcc gagagccca agggccctgac cccggccctg agctggctgc tgcagaggat 480
 gcaggcccg gtcggggcc agaccccaaa gacggcctgc gtatgatcc ctcagctgtc 540
 ggaggaggcg gtggggggcg gggctccggg cagtgcgtgc ggcgtctgg ctgcctgtc 600
 ggaccatgtc aggagcgggtt cttgcttcca cgccttgcgg agccctcagt acttcgtgaa 660
 ctttgcatttgc cagcagcaca gcagcggaggctt ccctatgcgtc ctggccgagc tgcagcgtt 720
 gatgcagcgc ctgggggtgg gcaggaggc ccacagtgcgtac cacagtcatc ggcacagggg 780

agccagcagc cgggaccctg tgcccctcat cagctccagc aacagctcca gtgtgtggga 840
 cacggtatgc ctgagtgcca gggacgtat ggctgcataat ggactgtcgaa acaggctgg 900
 ggtgaccccg gaggccctggg cccaaactgag ccctgcctg ctccaaacagc agctgagtgg 960
 agcctgcacc tcccagtcca ggccccccgt ccaggaccag ctcaagccatg cagagaggta 1020
 tctgtacgc tcccgtggca cgctgctcat ctgcctctgc ggggtcttgc gcctcctgt 1080
 gctgacccgc actggctgc ggggggtcac ccactacatc ctgcagacatc tcttgaccc 1140
 ggcagtggtt gcaactactg gggacgtgt cctgcataatc acggccaaagg tgctgggct 1200
 gcatacacac agcgaagagg gcctcagccc acagccccc tggccctcc tggctatgtct 1260
 ggccgggctc tacgccttct tccctgttgc gaaaccttcc aatctccatc tgcccaggaa 1320
 cccggaggac ctggaggacg ggccctgcgg ccacagcagc catagccacg gggggccacag 1380
 ccacgggtgtc tccctgcagc tggcacccatc cgagctccgg cagcccaagg ccccccacga 1440
 gggctcccgcc gcaagacctgg tggcgaggaa gagcccccggg ctgctgaacc ctgagcccg 1500
 gagactgagc ccagagttga ggctactgccc ctatatgatc actctggggcg acgcccgtgc 1560
 caacttcgccc gacgggctgg ccgtggggcgcc cgccttcgg tccctcttgc agaccgggt 1620
 ggccacccctg ctggccgtgt tctgcccacga gttccacac gagctggggg acttcggccgc 1680
 ctgtgcac gcggggctgt ccgtgcgcac agcaactgtg ctgaacccctgg cctccgcgt 1740
 cacggccctc gctggctctc acgtggcaact cgcgggttgg gtcagcgagg agagcgaggc 1800
 ctggatccctg gcagtgccca cccggctgttgc cctctacgtt gcaactctgcg acatgtccc 1860
 ggcgatgttgc aaagtacggg acccgccggcc ctggcccttc tccctgtgc acaacgtggg 1920
 cctgctgggc ggctggaccg tccctgtgt gctgtccctg tacgaggatg acatcacctt 1980
 ctgataccctt gcccttagtcc cccaccccttgc acttaagatc ccacacctca caaacctaca 2040
 gcccagaaac cagaacccccc tatagaggcc ccagtccttca ctccagtaaa gacactcttgc 2100
 tccttggaaa aaaaaaa 2117

<210> 54
 <211> 1495
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1452856CB1

<400> 54
 ctgaaatccc gcgaggatca acccgagctcg ccgaaaggag ggaggaacgt atcccttctg 60
 gaggctgtct cagggggcag agggacccga ccggaatgtc cgtgagcggtt ttccgggtt 120
 ctggagccca gccgggggttgc ttagagtcgg taaggagcag ctccaggat cctgagatcc 180
 ggagcagccg gggtcggagc ggctcctcaa gagttactgtc tctatgaaat ggcagagaat 240
 ggaaaaaaattt gtgaccagag acgtgttagca atgaacaagg aacatcataaa tggaaatttc 300
 acagaccctt ctcaagtgtaa tggaaaggaaagg aggaggagc gggaaaggaaag gcagaatattt 360
 gtcctgttgc gacagccgctt cattacccatc cagtatccatc ctctggaaat ccttgcata 420
 ttgaaggaat ggacctcaaa attatggcat cgtcaagca ttgtgggttc tttttactg 480
 ctgttgcgt tgcttatagc tacgttattat gttgaaggag tgcatcaaca gtatgtgca 540
 cgtatagaga aacagtttctt ttgtatgtcc tactgttagat gcttaggaat ttgttctt 600
 gttgggttgc gAACAGGGCTT gcacacccatc ctgtttatc tgggtccaca tatagccctca 660
 gttacatattatc ctgtttatgtca atgcaatttca gttatatttcc ccaaccacc ctatccgtat 720
 cagattattt gtcctatgtca agagggtactt gaaaggaaatcc tttttttgtt ggtatcattc 780
 tcaaaaatgtt ggtatgttgc ctgcattgtt ggtatgttgc cagcaatcg agagctgcct 840
 ccatatttca tggccagagc agctgccttc tcaggtgttgc aaccagatgtc tgaagagat 900
 caggaatttgc aagagatgtt ggaacatgtca ggttgcac aagactttgc ctccggggcc 960
 aaactggcag ttcaaaaactt agtacagaaa gttggatttt ttggaaatttt ggcctgtgt 1020
 tcaattccaa atcccttattt tgatctggctt ggaataacgt gtggacactt tctggatcc 1080
 ttttggaccc tctttgggtgc aaccctaaat gggaaaagcaa taataaaaaat gcatatccat 1140
 aaaattttgc ttataataac attcagcaag cacatgttgc agcaaatgtt ggttttccatt 1200
 ggtgctgtcc ccggcatagg tccatctctg cagaagccat ttcaggatgtt cctggaggct 1260
 caacggcaga agcttcacca caaaaggcga atgggcacac cacaggagaa aaactgggtt 1320
 tcctggatgtt ttggaaaatgtt ggtcgttgc atgggtgtt acttcatccat atctatccatt 1380
 aactccatgg cacaaggatgtt tgccaaacgtt atccagcagc ggttgaactc agaggagaaa 1440
 actaaataag tagagaaaatgtt tttaaactgtc agaaatttggg gttggatgggt tctgc 1495

<210> 55
 <211> 1747
 <212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1562471CB1

<400> 55

accagcagaa ggctgggagt ctgtagttt ttcctgctgc caggctccac tgagggaaac 60
 ggggacctgt ctgaagagaa gatgccccctg ctgacactct acctgctcct cttctggctc 120
 tcaggctact ccattgccac tcaaatacacc ggtccaaacaa cagtgaatgg cttggagcgg 180
 ggctccttga ccgtgcagtg tgtttacaga tcaggctggg agacctactt gaagtgggtt 240
 ttcgaggag ctatggcg tgactgcaag atcctgtta aaaccagtgg gtcagagcag 300
 gaggtgaaga gggaccgggt gtccatcaag gacaatcaga aaaaccgcac gttcactgtg 360
 accatggagg atctcatgaa aactgatgt gacacttact ggtgtggaa tgagaaaact 420
 gaaaaatgacc ttgggttac agttcaagt accattgacc cagcaccagt caccctaa 480
 gaaactagca gctccccaaac tctgaccggc caccacttgg acaacaggca caagctcctg 540
 aagctcagtg tcctcctgcc cctcatctc accatattgc tgcgtcttt ggtggccggcc 600
 tcactcttgg cttggaggat gatgaagtac cagcagaaag cagccgggat gtccccagag 660
 caggtactgc agcccttggg gggcgaccctc tgctatgcag acctgaccct gcagctggcc 720
 ggaacctccc cgaaaaaggc taccacgaa ctttcctctg cccagggttga ccaggtggaa 780
 gtggaatatg tcaccatggc ttcccttgcgg aaggaggaca ttccctatgc atctctgacc 840
 ttgggtgtc aggatcaggaa accgacactc tgcaacatgg gccacccatcg tagccaccc 900
 cccggcaggg gcccctgagga gcccacggaa tacagccaca tcagcaggcc tttagctgca 960
 ctccaggctc ttcttggac cccaggctgt gagcacactc ctgcctcatac gaccgtctgc 1020
 cccctgcctc cctcatcagg accaaccggg ggactggcgc ctctgcctga tcagccagca 1080
 ttggcccttag ctctgggttgc ggcttggggc caagtcctcg ggggcttcta ggagttgggg 1140
 ttttctaaac gtcccttcct ctcctacata gttgaggagg gggctaggaa tatgctctgg 1200
 ggcttcatg ggaatgatga agatgataat gagaaaaatg ttatcattat tatcatgaag 1260
 taccattatc ataatacaat gaacctttat ttattgccta ccacatgttga tgggctgaat 1320
 aatggccccc aaagatatct gtgtcctaatt cctcagaact tgcgtactgtt accttctgtg 1380
 gcagaaaaaggc acagtgcaga tgcgtatgtaa ttaaggactt tgagatagag aggttattct 1440
 tgctgtatca ggtggggccca aaatatcacc acaagggtcc tcataagaaa gaggccagaa 1500
 ggtcaaaagag gttagagacaa agtgcgtatgtt gaaaggacg tgggtgtgac gtgagcagg 1560
 gccatgaatg ccgcagccct catatggccag aaaggaaag gaatggattc ccctgcctgg 1620
 agcctccaaa agaaaccacg cctgcccacg ctttgcatttgc agcccatttgc aactgatctt 1680
 gagctcttgg cctccagaat tgcaggagaa taaatttgc ttgtttttaa tgaaaaaaa 1740
 aaaaaaaaaa 1747

<210> 56

<211> 1473

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1618158CB1

<400> 56

acagcgggac agccggcag ctttgggctt ccagcgctgg cttcatcgcc acctccctgc 60
 cttctccac ttcccgattt ttaggtgacg gatggggca ccatcaagca aaagatctt 120
 accttcgacg ccatgttctc caccacatcg tcaacatgg agaactaccg caagcgagag 180
 gacctgggttgc accagtccac tgcgtggctt cccgaggctcc ggatctcaga caatgggtccc 240
 tatgagtgcc atgtgggcat ctacgaccgc gcccaccagg agaagggtggt cctggcatca 300
 ggcaacatct tcctcaacgt catggctctt cccaccttca ttgaagtgggt ggctgtgtac 360
 acaccagccc ctttcagccg ctaccaagcc cagaacttca cgctggcttgc catcggtct 420
 ggagggaaaac cagcaccat gtttatttc aaacgagatg gggaaaccaat cgacgcgtg 480
 cccctatccatc agccaccacg tgcgtggctt ggcccccctac aggacacggc gcccctccgc 540
 agccttcgtc accgtgaccc ggtgcacacc aagatgcaga agtgcactgtc cttccctggac 600
 gcccggagaacc ggggtggggc accctacacg gagcggccctt cccgtggcc gaccccgat 660
 cccaaacatcc tcctccagcc aaccatccatc agacgggtcg gggccgtgag 720
 ttcccccgtt gggccacacg cggccggcc accctacttcc tgcggccacag ccgcaccccg 780
 agcagtgcacg gactgtggaa agtgcgttcc ctgcgtcacctt gaccccttcaaa cccacacatc 840
 gacaacgaggcc cctcttcag ctgcgtggcc aagcaccatc gcccgtgtacat gcccacatc 900

gcagaggta cgctgggtgc ccccaaagga cccaaaattg ttagtgcgc cagcagagcc 960
 cgggttagggg acacagttag gattctgtc catgggttc agaacgaagt cttccggag 1020
 cccatgttca cgtggacgcg ggttgggagc cgcctcctgg acggcagcgc tgagtgcac 1080
 gggaaaggagc tgggtctgga ggggtttcc gccgagctca atggctccat gtatcgctc 1140
 accgcggcacc acccaactggg ctccaccggac acgcacacccc ggctcatcg tttgaaaac 1200
 ccaaatatcc caagaggaac ggaggactt aatggttcca ttggcccccac tggtgcggg 1260
 ctccatctgg tgcgtccct gacagtatt ctggagctga cgtgaaggca cccggccgg 1320
 ccactccatc aggcaactgac atctccacga cgggtttca ttcttttct aaactatttc 1380
 cagtctgtt cttagtctt ttccatctgt gtcttggctt cttcagtccg ttaattaaa 1440
 acaaacagaa caattttccc caaaaaaaaaaaa aaa 1473

<210> 57
 <211> 1591
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1656935CB1

<400> 57
 attaaacacac tcactatagg gaatttggcc ctggaggcaa gaattcggca cgaggccttc 60
 tataggagag ctttagaaatg aagtcatatc gaggaggcagg cgaatcctaa ccgcattctct 120
 ctcttagct ggactgaacc caaacatgaa tgtaacacgc atggacatga ccgggtggctt 180
 gtcggtaag gaccatccc agtcccaagtc acgcctcccc cagtggacgc accccaactc 240
 catggataac ttggccatgt ccgttcccc cctggagcag aacccttagca agcatggtgc 300
 tatccctgga ggtctaagca ttggccctcc aggttaagtcc tccattgtat actcctatgg 360
 ccggtagat ttaatccaga acagttagtc accagccagt cctccctgt ctgtccccca 420
 tagctggta cgtgccaaat ctgacagtga taaaatctca aatggctcta gcatcaactg 480
 gccccccagaa ttccatccgg gagttccatg gaaaggactg cagaatattg accctgagaa 540
 tgaccctgac gtcactcctg gcagtgccc cactggcctt accatcaaca ccaccatcca 600
 ggtatgtcaac cgctacccctc tcaagagtgg agggtcctcc cggccatcat ctcagaatgc 660
 cacgtcgctt cttcgagtg cctggccact cagtgcctt ggttccatg gcttttca 720
 cagcattgca tccgcaccta gtgttgcagg taaactgtca gacatcaaat cgacgtggc 780
 ctctggccctt acctccacca cgcacccctc tctgtctcat gaactatgg aggtgcccag 840
 aaacagtaact gcacccacga ggcacccctc aggttaacc aatcccaagc cctccctccac 900
 ctgggggtgcc agcccccctcg gctggaccag ctccctactcc tcgggttctg cctggagcac 960
 cgacacctca ggaagaacca gcagctggct cgttcttgc aacctcaact cccagattga 1020
 tggttctaaa ctgcggacat tgggttgc acatgggcctt cttatcacat tccacctgaa 1080
 tctgactcaa ggcacatgtg tggccgtt cagtcaccaag gaggaggct tgccaaaggc 1140
 ccaggaagtc cttgcaccaa ttgtgcgtcc ttgggagaca cttagccatt cttggggccc 1200
 gagttccgc ttgggtggaa cgaaggaaat tggaaattcgg gtttccat tggccaccc 1260
 gggcccccaggc cgcattggcc agagcacaat ttccaaaggc ctggccacat tccacgacca 1320
 gagggggcgtt tccaagctca caggacgtgg cgggattcac aggccacgcg ggagggggca 1380
 ggcttctcac caacttcac atatggaca ttgtgaatta acattttaag agcgttgcga 1440
 ctaatggga caatcatcgga acaaggaccc tggatcaat ttgtaaatg gcccctttat 1500
 aaatgttgc cgtgttagaat aaaggttctt ttccactata tcccatgttgc tcggatata 1560
 tcacatcgta acagtgttat ttgggaa a 1591

<210> 58
 <211> 1858
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1859305CB1

<400> 58
 gcagggcgccg cggaaacttcc ctctacccgc cggcccccgcg gcgcgcaccc ttggcgctgg 60
 acgcttcctc cttggaaagcg cctctccctc aggtctcaag aactgtttt agatgcagga 120
 attcatctaa ttgttactgc cggccgagggt gtgagagccc tagcatctga aagtggcga 180
 cttggaggtt gttatggaga aaacttggat agatgcactt ccttacta tgaattcttc 240

agaaaagcaa	gagactgtat	gtatttttg	aactggtgat	tttggaaagat	cactgggatt	300
gaaaatgctc	cagtgtggtt	attctgttgt	tttggaaagt	cgaaacccccc	agaagaccac	360
cctactgcc	agtgggtcag	aagtcttgag	ctattcagaa	gcagccaaga	agtctgacat	420
cataatcata	gcaatccaca	gagagcatta	tgattttctc	acagaattaa	ctgaggttct	480
caatggaaaa	atattggtag	acatcagcaa	caacctcaa	atcaatcaat	atccagaatc	540
taatgcagag	tacccctgctc	atttggtgcc	aggagcccac	gtggtaaaag	catttaacac	600
catctcagcc	tgggctctcc	agtcaaggagc	actggatgca	agtccggcagg	tgtttgtgt	660
tggaaatgac	agcaaagcca	agcaaagagt	gatggatatt	gttcgttaatc	tggacttac	720
tccaaatggat	caaggatcac	tcatggcage	caaaggaaatt	gaaaagtacc	ccctgcagct	780
atttccaatg	tggaggttcc	ccttcttatt	gtctgtgtg	ctgtgtgtc	tctttgtttt	840
ctattgttt	ataagagacg	taatctaccc	ttatgtttat	gaaaagaaag	ataatacatt	900
tcgtatggct	atttcatttc	caaatcgat	ctttccaata	acagcactta	cactgcttgc	960
tttggtttac	ctccctggtg	ttattgtcgc	catttctacaa	ctgtaccgag	gcacaaaata	1020
ccgtcgattc	ccagactggc	ttgaccactg	gatgtttgc	cgaaaggcgc	ttggcttgg	1080
agctctggga	tttgccttcc	ttcatgttct	ctacacactt	gtgattccct	ttcgatattt	1140
tgtacgatgg	agattgggaa	acttaaccgt	taccaggca	atactcaaga	aggagaatcc	1200
attttagacc	tcctcagcc	ggctcagtg	ttcatatgtg	gctttggaa	tacttgggtt	1260
ttttctgttt	gtactcttgg	gaatcactc	tttgcacatc	gttagcaatg	cagtcactg	1320
gagagagttc	cgatttgtcc	agtccaaact	gggttatttg	accctgatct	tgtgtacagc	1380
ccacaccctg	gtgtacggtg	ggaagagatt	cctcagccct	tcaaaatctca	gatgttatct	1440
tcctcagcc	tacgtgttag	ggcttacatc	tccttgact	gtgctggtga	tcaagttgt	1500
cctaatacatg	ccatgtgtag	acaacaccc	tacaaggatc	cgccagggt	gggaaaggaa	1560
ctcaaaacac	tagaaaaaagc	attgaatgg	aaatcaat	ttgaaacaaa	gttcaattta	1620
gctggatttc	tgaactatgg	tttgaatgt	ttaaagaaga	atgatgggt	cagtaggaa	1680
agtttttttc	ttacaccgtg	actgagggaa	acattgttg	tctttgagaa	attgactgac	1740
atactgaaag	agaacaccat	tttatctcag	gttagtgaag	aatcagtgc	ggtcctgac	1800
tcttattttc	ccagaggcca	tggagcttag	attgagacta	gccttgggt	tttcacac	1858

<210> 59
<211> 1454
<212> DNA
<213> *Homo sapiens*

<220>
<221> misc_feature
<223> Incyte ID No: 1949083CB1

<400> 59	caccctgggt	tctccctgt	tggccctgca	aggcatatca	tgcttcctgt	ctctaggacc	60
	tgcttactag	aaagcagcac	ccgcctgaaa	cctcacgaag	cccagaacta	caggaagaag	120
	gcatttgggg	tgtcctgggt	ctccatcatt	gtcaccctgg	ccctcgcggt	ggctgcctt	180
	actgtctccg	ttatgaggtt	cagcgcctct	gcttttgggt	ttgcatttga	tgccatctgt	240
	gacgtctgt	catcggcgt	tgtcctgtgg	cgttacagca	acgcggccgc	tgtgcactct	300
	gccccataggg	agtacatagc	ctgtgtcata	ttgggggtga	tatttcctt	gtcatccata	360
	tgtatagtgg	tcaaaagccat	ccatgaccc	tcaactaggg	tgctcccaga	agtggacgat	420
	ttccctgttca	gtgtctccat	ttaagtggg	atttggca	gcatcctggc	cgtgttgaag	480
	ttcatgttgg	ggaagggtct	gaccagtata	gcactcataa	cagatgggtt	taactccctc	540
	tggggtggcg	tgtatgggctt	ctccattttt	ctgagcgcgg	aagtgttcaa	gcatgactcg	600
	gcgggtctgg	acctggacgg	cagcataggg	gttctgatcg	gcctcaccat	atttgcctat	660
	ggggtcaaaac	tcctcatcga	catggtggcg	aagggtgaggg	agacacgtca	ctacagatgt	720
	ttttagtggaa	ggggggccagc	atccgcata	gaccatcgag	atgaggagtt	cccacatagg	780
	caaagggtgc	caatatttaa	ctgaacatct	ggtttctttt	tggaaagttt	ctttcacatg	840
	gtttgtcatt	acaagacaag	gtctgcccag	ccaggtggat	ctacccctgoc	cccatcacct	900
	gccgccccca	tcaaaacatgt	tgggacaatg	cccatagggaa	tggacctct	tccccgtctc	960
	cagctgggac	tggcggtttt	tagtctctgg	agtatgtgg	ttctctatggg	taggtatgaga	1020
	tctttggcag	aaaggcttcc	gggtgggtgtc	tgagcctgcg	ctgcataatgg	ctgagcagac	1080
	ccaccccttc	cagcttgggt	ggccctggca	ctcctgggtc	caagtcttc	ctttcctggc	1140
	aggctctaag	ggaaagattgt	acccttcacc	ctttacatac	ccagaatcat	cagtatgtca	1200
	cttcttaatt	tctatcagt	tatcttcatt	tttcatactg	ttttactaat	cctaagtctt	1260
	aacagatttgc	ctcaaaaaggat	gaccatttca	ttttttaaag	tacttagtga	tacacgtata	1320
	agctttgcatt	ggacgaatttgc	aataagcaca	ttgacctttt	cttgcacatt	cagaacctgaa	1380
	acatccatgt	gaaaactggg	tccattttg	agagatgtga	aactacagtt	tatatgttaat	1440
	aaataaaat	aata					1454

<210> 60
 <211> 2310
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1996357CB1

<400> 60
 ggcccgatg ttcgggtcag ctgccagatc cgctgatcta gtgcttctcg aaaaaaacct 60
 tcaggcggcc catggctgtc gatattcaac cagcatgcct tggactttat tggggaga 120
 ccctattatt taaaatggc tcaactgaaa tatatggaga atgtggggta tgcccaagag 180
 gacagagaac gaatgcacag aaatattgtc agcctgcac agaatctcc 240
 atggctcta tcttgattt atggcaatgc ttccctctgg tttacattgg ttcttcattg 300
 aatggtactc gggaaaaaag agtccagcg cactttcca acacatact gcattattt 360
 aatgcagcat ggcagctatt atcaccttac ttgtgagtga tccagttgg gttcttata 420
 ttcgttcat tgcgttattt atgtttctg actgtacac gatgtttac aacccaagtc 480
 cagattacgt taccacagta cactgtactc atgaagccgt ctacccacta tataccattg 540
 tatttatcta ttacgcattc tgcttggat taatgtatgc gctccgaccc ttcttggta 600
 agaagattgc atgtgggta gggaaatctg atcgattaa aagtatttat gctgcactt 660
 acttcttccc aatttaacc gtgcttcagg cagttggtgg aggcccttta tattacgcct 720
 tcccatacat tatattatgt ttatctttt ttactctggc tggatcatgt tctgcttctg 780
 aaatagagaa ctgctatgtat ctctggcata gaaagaaaag acttattgtt ctcttcagcc 840
 actggtaact tcatgcctat ggaataatct ccatttccag agtggataaa cttgagcaag 900
 atttgcctt tttggctttt gtacctacac cagccctttt ttacttggc actgcaaaat 960
 ttaccgaacc ttcaaggata ctctcagaag gagccaatgg acactgagtg tagacatgt 1020
 aaatgcacaa aacctgagaa gtgctcctaa taaaaaagta aatcaatctt aacagtgtat 1080
 gagaactatt ctatccatata tggaaacaag attgtcaga tatcttaatg tttgggttt 1140
 tctttgtttt gttatggtt agacttacag acttggaaaa tgcaaaaactc tgaataactc 1200
 tggatcacag ggtatataat tctgtacac tggaggccg ctaggaagcc cttgcttctc 1260
 tcaacagttc agctgttctt tagggcaaaaat tcatgtttt gtgtacctt caatgtgtc 1320
 ccattttattt aagaaaagct ttaacacgtg taatctgcag tccttaacag tggcgtatt 1380
 gtacgtaccc tttgtgtttt agttgtttt tcacccataa tgaattgtaa aaacaaacat 1440
 acttggggg tctgatagca aacatagaaaa tggatgtat tttttttt tttttttt 1500
 ttttcatcaa tacagtattt tggatgtattt caaaaataga taataattta tataacaggt 1560
 tttctgtttt tagattgggtt caagattttt tggattattt gttccctgtaa agaaaacaat 1620
 aataaaaagc ttacccatata aaaaatttcaaa tggatgtttt tttttttt 1680
 atagtatgga agtaattcaa actggtaat agtttccctt catatctcg gtatataac 1740
 ataccatatt ttattgtatcc agagataactt atttccatctt gtgacatctc tgaatttagga 1800
 tgcacatcttac aactgtatggc ttatttaggtt taatggaaa cagaagatatac acagtataaa 1860
 aagggttttctgtgtttt tttgtgtttt gtgtatgtt tttatgtttt 1920
 gaaggcccta agactcatgg ttgcaaccat ggaagcaaaa tgaaatttt agctcttaac 1980
 ctaacaaccc gaccatgttcc atccatttttt attgtttaga agttttttt ctgataactt 2040
 gtggagggtt tggatgttcc taaaattttt aatgtttagt accttcttata acagctgcaa 2100
 aatatgaaag taagtgcact cactttccctt gtgtatgtt tttttttt 2160
 ttgtatccctt gagttactttt gttatgtat tttttcttactt acattttacc actggaaat 2220
 gaacccttgc acgaatgtgt ttcttcttctt ctgttaggaat aaaaaataaa tataaaaattt 2280
 ttatgtat tgcacacaaa aaaaaaaaaaaa 2310

<210> 61
 <211> 744
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2061330CB1

<400> 61
 agtgggtgt tacctgcccga cagcataaaag cgaggcaagg tccagccgtt ccgccccggca 60
 tacaagctt tggaggcagcc ccctttaag atctccaggt ggacttcaca gagatgtcaa 120
 agtgttagagg tggatcaaga actggaaacgt agccttttgc tggatctgt 180

ggaaaggacc	ccagactgtc	gttctgagca	ctccccaccgc	tgtgaaggta	gaaggaatcc	240
cagcctgat	ccaccacagc	catgtaaaac	ctgcagcgcc	tgaaacctgg	gaggcaagac	300
caagcccaga	caacccctgc	agagtgaccc	tgaagaagac	gacaagccct	gctccagtca	360
cacccggaag	ctgactggc	cacgcacggc	cgaagcctga	ggaagctcat	catgagattc	420
atttttctta	aattttggac	ttatacagta	agggcttcaa	ctgatcttac	tcaaactggg	480
gactgttccc	agtgtactca	tcaggtcacc	gaagtagggac	agcaaattaa	aacaatcttt	540
ctgttctata	gttattatga	atgtatggaa	acaataaaag	aaacttgtt	gtataatgcc	600
actcagtgaca	aggtatgtag	cccgagaaaat	gaccgacctg	atgtgtgtta	taacccatct	660
gagccccctg	caccaccgtt	tttggaaataa	gaataagaac	tggcctttc	ctaggtgata	720
caaqtaaaat	aataactaga	acag				744

```
<210> 62
<211> 1109
<212> DNA
<213> Homo sapien
```

<220>
<221> misc_feature
<223> Incyte ID No: 2346947CB1

<220>
<221> unsure
<222> 30
<223> a, t, c, g, or other

<400> 62	gaagcagtgc	agagaggaga	gcggagcgcn	agtgccgctg	agcaaaggcc	ttcaccatgg	60
ccgagtcccc	cggctgctgc	tccgtctgg	cccgtgcct	ccactgcctg	tatagctgcc	120	
actggaggaa	atgccccaga	gagaggatgc	aaaccagcaa	gtgcgactgt	atctggttt	180	
gcctgcttt	cctcaccttc	ctcccttccc	tgagctggct	gtacatcggg	ctcgcccttc	240	
tcaatgacct	gcacaacttc	aatgaattcc	tcttcggccg	ctggggacac	tggatggact	300	
ggtccctggc	atttctgtcg	gtcatcttc	tactggtcac	atatgcatcc	ttgcatttgg	360	
tcctggccct	gctcctgccc	cttggtagac	agccccgtca	tctgcacagc	ctccacaagg	420	
tgctgtctgt	cctcattatg	ctgcttg	cggctggct	tgtggactg	gacatccaat	480	
ggcagcagga	gtggcatagc	ttgcgtgtgt	cactgcaggc	cacagcccca	ttccttcata	540	
ttggagcagc	cgtctggatt	gcccctctgg	cctggcctgt	ggctgatacc	tttctaccgt	600	
tccaccgaag	aggccccaaag	attctgtctac	tgctcctatt	ttttggagtt	gtcctggtca	660	
tctacttggc	ccccctatgc	atctcctc	cctgcatcat	ggaacccaga	gacttaccac	720	
ccaagcctgg	gctgggtggga	caccgaggg	ccccatgt	ggctcccgag	aacaccctga	780	
tgtcctgct	gaagacagct	gaatgcgac	tactgtgtt	gagactgtat	tgatggtcag	840	
ctccgtatgg	gtcccccttcc	tcatgcata	tgagcacctc	agcaggacca	cgaatgttagc	900	
ctctgtattc	ccaaacccgaa	tcacagccca	cagcagtgc	ttctcctgga	ctgaactgaa	960	
gagactcaat	gctggatcct	ggttcctaga	gaggcgaccc	ttctgggggg	ccaaaccgct	1020	
ggcaggccct	gatcagaaaag	aggctgagag	tcagtcggta	ccagcattag	aagagctatt	1080	
ggaggaagct	gcagccctca	accctttca				1109	

<210> 63
<211> 2511
<212> DNA
<213> *Homo sapiens*

<220>
<221> misc_feature
<223> Incyte ID No: 2795577CB1

```

<400> 63
gcagccgtcg cttcggagc gaagggtacc gaccggcag aagctcgag ctctcggggt 60
atcgaggagg caggcccgcg ggcgcacggg cgagcgggcc gggagccgga gcccgggagg 120
agccggcagc agcggccgccc cgggtccacg gcgaggccgt cgacgtctct gaaaactgc 180
gcccggctc gcccgtactgc gcccggagcg atgaagatgg tccggccctt gacggcggtt 240
tactccaaaca gctgtctttt gtgtccat gtccgcaccc gcaccatctt gctcgccgtc 300
tgttatctgtatc tcatcaatgc ttgtgtactg ttgattttat tgagtgcctt ggctgatccg 360
gatcaatatac actttcaag ttctgaactg ggagggtactt ttgagttcat ggatgatgcc 420

```

aacatgtgca ttgccattgc gatttcttctt ctcatgatcc tgatatgtgc tatggctact 480
 tacggagcgt acaagcaacg cgcagcctgg atcatccat tcttctgtta ccagatctt 540
 gactttggcc tgaacatgtt ggttgcatac actgtgctt tttatccaaa ctccattcag 600
 gaatacatac ggcaactgcc tcctaaatcc tcttgcatac atgatgtcat gtcagtgaat 660
 cctacctgtt tggcccttat tattttctg tttattagca ttatcttgac ttttaagggt 720
 tacttggat tctgtgttg gaaactgcatac cgatacatca atggtaggaa ctccctgtat 780
 gtcctgggtt attttaccat gaaactgcatac acgggtctgc taccgggtt tgatgtgcc 840
 actgtgaatg gtgcgtccaa ggagccaccc ccacccatcg tgcgtgcatac agcctcaag 900
 tggccggagc tgagggcagc agcttgactt tgccacatca tgagcaatag ttctgttatt 960
 tcactttgc catgagccctc tctgagctt tttttgttgc aatgtactt ttttaaaatt 1020
 tagatgttag attgaaaact gtagtttca acatgttgc ttgttgcatac ctgtgataga 1080
 ttaactgttag aatttttctt gtacgattgg ggtatataatg ggcttcacta accttcccta 1140
 ggcattgaaa ctccccccaa atctgtatgg cctagaatgc tgctttgtt cctgtgggc 1200
 cccaaagggtt ggcatttttc tctctgttcc ctctctttt aaaaatgtaaa ataaaaccaa 1260
 aaatagacaa cttttcttc agccatttcca gcatagagaa caaaacctt tggaaacagg 1320
 aatgtcaatt gtgtatcat ttttcttattt aggtttatag aagtccctt gtagtgcgtt 1380
 caagaatttc ccccacacaaca tcctttatgtt ctgaaggatca atgacatgtt gtgttgggt 1440
 gtaaaaggatt ttccatgg cctgttattt gaccattttt aagcaccagg ccgtgggagc 1500
 agtgaccatc tgctgtactt tttttgttgc tttgttgcatac gggacatggg gtgacatgcc 1560
 tcgtatgtt tagagggtgg aatggatgtt tttttgttgc tttttgttgc tttttgttgc 1620
 ttctccttgc ttcacatccc caccgggc ccgttttac taatgtttt gcccctagatt 1680
 ggttcaagga ggtcatccaa ctgttttttcaatgttgcatac tttttgttgc tttttgttgc 1740
 tctgcctaac aacatggaaa agggtttttctt tttttgttgc tttttgttgc tttttgttgc 1800
 aacttccaaat tatgtctatg ttttttttcaatgttgcatac tttttgttgc tttttgttgc 1860
 gccttccttgc tatgttgc ttttttttcaatgttgcatac tttttgttgc tttttgttgc 1920
 catatgttac aaagtccatgc ttttttttcaatgttgcatac tttttgttgc tttttgttgc 1980
 gtcgtttgc ttttttttcaatgttgcatac tttttgttgc tttttgttgc tttttgttgc 2040
 aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa 2100
 aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa 2160
 aaatccctcgatgtgcataatg aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa 2220
 gtcgggggttgc ttttttttcaatgttgcatac tttttgttgc tttttgttgc tttttgttgc 2280
 ctctttttaatgc ttttttttcaatgttgcatac tttttgttgc tttttgttgc tttttgttgc 2340
 atatatctcccttgc ttttttttcaatgttgcatac tttttgttgc tttttgttgc tttttgttgc 2400
 cccgcgccttgc ttttttttcaatgttgcatac tttttgttgc tttttgttgc tttttgttgc 2460
 ccagcggcgggggttgc ttttttttcaatgttgcatac tttttgttgc tttttgttgc tttttgttgc 2511

<210> 64
 <211> 788
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3255825CB1

<400> 64
 cccacgcgtc cgcagactgc tcctctggc caccgtccac ttggccgcctt ccacccggaa 60
 agccccccat gctgatgtgc gcatgtatcc catcaccgaa tggcagaaga ttgggttggg 120
 gatcaccgggtt ttcggcatct tcttcatctt ctttggaaaca ctccctgtact ttgattccgt 180
 gtcctgttgc tttggaaacc ttgttgcatac gacgggcctt tccctcatca ttggccttgc 240
 gaagaccttt ttgttgcatac tccaaacggca caaactcaag ggaaccagct tcctcttgc 300
 ggggtgtgtt atcgtgttgc ttttttttccatgttgcatac tttttgttgc tttttgttgc 360
 attcttcaggc ctctttaagg gcttttttccatgttgcatac tttttgttgc tttttgttgc 420
 aacatcccttgc ttttttttccatgttgcatac tttttgttgc tttttgttgc tttttgttgc 480
 aaaaacagaga ttgttgcatac ttttttttccatgttgcatac tttttgttgc tttttgttgc 540
 gaaccacccc ctcgttccatgttgcatac ttttttttccatgttgcatac tttttgttgc tttttgttgc 600
 tccagaagga aggaatggag ctgagcaact gacgtcaat ccccaagtgc actcaagagg 660
 ctgcccggaa gcagatgtgc agaccccaag gagactggc tggggcttgc atcacaccc 720
 cactcttatat ttatgggagg aaaaatgttgc ttttttttccatgttgcatac tttttgttgc 780
 aaaaaaaaaaaaaaaa 788

<210> 65
 <211> 1831

<212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3393430CB1

<400> 65
 gccttataaa gtagcctctg catctgcctg ctcggccag aggagggcta ccctgggct 60
 gagagttcac ctgtctcagg aaccacatga gcccacatg cctgtggca gcggccagg 120
 cagccatggc ttggcaagt aggctggcc tgctgctggc actgctgctg cccgtggcg 180
 gtgcctccac gccaggcacc gtggccgac tcaacaaggc agcattgagc tacgtgtctg 240
 aaattggaa agccctctc cagcggggcc tgcaggatc tgcaggatcat ttccctggact 300
 ggagtggaga ggcgttcag cccaccatg tccggattct gaatgtccat gtgcccccc 360
 tccacctgaa attcattgtc gtttcggag tgcgcctgct ggcagcagct aattttactt 420
 tcaaggcttt tcggccccca gagccccctgg agctgacgct gcctgtggaa ctgtggctg 480
 acaccgcgt gaccaggagc tccatcgaga cccctgtggt cagcatctct gcctgtctt 540
 tattctcggg ccacgccaac gagtttgcgt gcaagtaacag cacctccac ggcgtgtgg 600
 tccctggcgca gaagcacatt aaagctgtct tgagtaacaa gctgtgcctg agcatctcca 660
 acctgggtc ggggtcaat gtccacatgg gcacccatgg tggcctcaac cccgtgggtc 720
 ctgagttcca gatccgttat tccatggtca gtgcgtccac tgcaccatg gactacattt 780
 ccctggaaat caatgcgttt ctcttcgttgc tggcaagcc catcatccgt cccacggatg 840
 ccacccctt tggatggcca aggcatgtgg gtaccgggg ctccatggcc accgtggcc 900
 tctccacgca gtcgtttgac tctgcgttcc tgctgtcgaa gaaggccggt gcccctaacc 960
 tggacatcac agggcagctg aggtcgatg acaacatctg gaacacatctt gctctggcc 1020
 ggctcatccc ggagggtggcc cgccagttt ccgagccat gcctgtgggt ctcaaggtgc 1080
 ggctgggtgc cacacatgtg gccatgtcc acacaaacaa cgccacccctg cggctgcagc 1140
 ctttcgttgc ggttctggcc acagcctcca actcggctt ccagtccttc ttctccctgg 1200
 atgtggtagt gaacttgaga ctccagctt ctgtgtccaa ggtgaagctt cagggacca 1260
 cgtctgtgtc gggggatgtc cagctcacgg tggcctccctc caacgtgggc ttcattgata 1320
 cagatcaggc ggcacactg atgggcacccg tttttgagaa gcccctgtg gaccatctca 1380
 atgctcttcc ggcacatggga attggccatcc ctggatgtggt caacccatccac tatgtcgccc 1440
 ctgagatctt tgcgtatgg ggctacgtgg tgatatccag tggactcttc taccagagct 1500
 gaggcaagac cactggggagg cctgagatg ggccagctcg ctgctcaggc gaatttctca 1560
 tttcaagcca ctggggaaac tgaggcaaaa ccatacttag tcatcaccaa caagctggac 1620
 tgcttagctg ggctgtttt tcttccttgc tgccctgggt ctccctccct cacttctgcc 1680
 ctttccttc ctccctccct tcttcctcat ctccctcat ctccctccct cttcctctgc 1740
 cccacccctg gggggagcag actgctccctc caggctgtat agacctgccc tcttcattt 1800
 aacaacttctt cttgagctgc aaaaaaaaaa a 1831

<210> 66
 <211> 1499
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3490990CB1

<400> 66
 tggggatgtt gggtaagaa ggccggaaatg taatccctttt ccaatcagct ccccatctcg 60
 ccctctggcg gtcggacc accggccgag ttactctgtat taccggccatg ctccaaagacc 120
 cgcactgaca gtccccacat cccgttctct gtcatcggtt ccctcaatgg agtccacatg 180
 tttggcaga atctggatgt gcagctgagc tctgcgagga ccgagaacac gactgtgggt 240
 tggaaaagct tccatgacag catcaccctc attgttctgt catctgaggt gggcatctct 300
 gagctgaggc tggagagact actccaaatg gtgtttggag ccatggctt tcttgggaa 360
 cttgaagaac tgaccaatatt ccccaacatg gagagactga agaaggactt gaggccagt 420
 tattgcctca tgcacatgtt cctggggacat tcggatgtca tcggggacat gaccaggatgt 480
 gtggactgcg tgatccctcc agaggggtcc ctcttcgggg aagcccttc cgggttcgtc 540
 gagggcccggtt gcaacgacat cgtcagttcg gtggatgtccg gcccgggtggt ggcaccaaca 600
 gaggggtgggtt ggcggctggg gacgccccatg gccgtgtcc tccctgggtt ggtgggggtcc 660
 ctggcccccgc agaccgctcg cgcactaccatc gtgtacatccat cgcacggggag cccacgggtc 720
 ccacacccggc tcctgacccat gactctgtcg ccgagccctgg agctgtgtctt actctgcggg 780

ccgagccac ccctcagcca gttgtatcca cagcttctgg agcgctggg gca
 ctggaccctg tgccggcctg tctgcccgtt ggacccccc cgctgcccag tggctcccc 900
 cttcacacag acatcctcg gctgctgctc ctccacctgg aactgaagcg ctgcctctc 960
 accgtggagc ccttggggga taaagagcct tcaccagaac agcgccggc cctccctccga 1020
 aacttctata ccctggtac c tccacgcac ttcccaccag agccagggcc accagagaag 1080
 acagaagatg aggtctacca gcccagctg cccagagctt gctacctggt gttggggact 1140
 gaggaaaggc gcacaggat gctgctggg gcctgcagc tggggcttcg gcccgtc 1200
 ctgctgtgt ctccccagat tccacccat gggctgcgaa gcctggccac ccacactctg 1260
 catgcctca cccacttct tgactacct agcagtgggt gatggacaca gacatgggc 1320
 tggtagcgtc tctctgtta ttcgctcaca ataatacaca gcccctggat ggggaggggg 1380
 taggaggggc tacaacaggg tgggggtggg ggggaggaga catccacttc cctggccct 1440
 ctccccctcg tgcttggggg gaaaggag ggaggggac tccccctaac ccccccagaa 1499

<210> 67
 <211> 365
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3635154CB1

<400> 67
 gtttttttgc actaacttca ggagccagct cgtgatctca ggatgttatgg aaaaataatc 60
 tttgtattac tattgtcaga aattgtgagc atatcagcat caagtaccac tggtgtggca 120
 atgcacacccat caaccttctc ttcaatcaca aagagttaca tctcatcaca gacaaatgg 180
 gaaacgggac aacttgtcca tcgtttcaat gtaccagctc ctgttagtcat aataactcatt 240
 attttgttg tggatggctgg tattatttggaa acgatcctct tattttctta cagtttcgc 300
 cgactgataa agggctgagg gtgtagcctg catgtgccc atcttgctct gaaccggccg 360
 cattg 365

<210> 68
 <211> 1102
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4374347CB1

<400> 68
 gaggacagag caggcagcag agaccatggg gccccctca gttgtcccc acagagaatg 60
 catccccctgg caggggctct tgctcacagc ctcacttta actttctggg acgcacccac 120
 cactgcctgg ctcttatttgc catcagcgcc ctttgaagtt gctgaagggg agaatgttca 180
 tctctctgtg gtttatctgc cggagaatct ttacagctat ggctggtaca aaggaaaac 240
 ggtggagggcc aaccagctaa tcgcagcatc tgtaatagac actcacgtt ggactccagg 300
 gcctgcatac aecgggtcggc agacaatata acccagtggg gatctgcatt tccagaacgt 360
 cacccttagag gacacggggat actacaaccc acaagtccaca tacagaaattt ctcagattga 420
 acaggcatct caccatctcc gtgtatacga gtcagtggtt cagccctccca tccaagccag 480
 cagcaccacca gtcacagaga agggctccgt ggtcctgacc tgccacacaa ataacactgg 540
 aacctcttc cagtggattt tcaacaacca gcgtctgcag gtcacgaaga ggatgaagct 600
 gtcctgggtt aaccatgtgc tcaccataga ccccatcagg caggaggacg ctggggagta 660
 tcagtgtag gtcctcaacc cagtcagctc caacaggagc gacccctca agctgactgt 720
 aaaatatgac aacactctag gcacccctgtat cggggctctg gttgggagtc ttctgtggc 780
 tgcacttgc tggatggatcc tccctccggaa aactggcagg gcccggatc agagtgactt 840
 cagggagccag cagcccccag cttccacccc cggccatggc ccctctgaca gctctgacag 900
 ctccatctcc taggaatttgc tacactctga cacaacatt tactgtggg tcgaccacaa 960
 agcagatgtg gtttcttagg ttccctctggg agctgctctt gtgggttgc ggagcgtccc 1020
 tgaaggcccccc agccctgggg atggggaaagg acatggagcc tgagccagag aaccagctct 1080
 gagtcctgag gagacacagg cc 1102

<210> 69
 <211> 2546

<212> DNA
<213> *Homo sapiens*

<220>
<221> misc_feature
<223> Incyte ID No: 4596747CB1

<400> 69

ccccggggcgc gccccaccgcg cccgcatccat gttcgacacc acacccccact ctggccggag 60
cacgccaagc agctccccat cgctccggaa acggctgcag ctccctgcccc caagccggcc 120
cccacctgag ccagaaccag gcaccatggt ggagaaggga tcagatagct cctcagagaaa 180
gggtggggtg cctgggaccc ccagcaccca gagccttaggc agccggaaact tcatccgcaa 240
cagcaagaag atgcagagct ggtacagttat gctgagcccc acttataagc agcgtaatga 300
ggacttccgg aaactgttca gcaaactccc cgaagcagaa cgcctcatgg tggattactc 360
ctgcgccccg cagcgtgaga tcctgctcca gggccgcctc tacctctctg agaaactggat 420
ctgcttctac agcaacatct tccgctggga gaccacgatc tccatccagc tgaaggaagt 480
gacatgtctg aagaaggaaa agacggccaa gctgatcccc aacgcctatc agatctgcac 540
ggagagcggag aagcattttct tcatttcctt tggggccctg gaccgctgtc tccttcctcat 600
cttccgcgtc tggcagaatg cactgttga aaagacgtg agtcccccg agctctggca 660
ccttggtgcat cagtgtctac gtcagagct gggcctcacc agtgaggatg aggactatgt 720
cttcccccctt cagctgaacg gtctggggac ccccaaggaa gtgggagatg tgatccctt 780
gagcgacatc acctccctgg gggcagctga ccgcagccag gagccaaggcc cagtggttcc 840
gcccgtggc catgtcacgc agaggaggac aaggaggagc gcaacccccc gcaacccccc 900
gacccgggtg gctgaaccccc aggtagacag ccagccagac gccttcctca gccagacagt 960
cgagcacaga gcccacccag cctgacggc ccaccacccct 1020
ccagtgagga gctattgaca gacacaagta actcccttcc 1080
acttggctgc cctgcttccc gacctctccg gcccctccct 1140
tgggcgttga gcccgtccag cagatgtctc tcteggaactc 1200
tacagcagtg caagttcaca gacgtgacgc tgagccccctg 1260
accagcggc ggtgtgacg tacaccatcc ccatacagaa 1320
cctccgtgtt ggagacacag acgctgttcc ggcggggccc 1380
tggactccga ggtgtgacg cagggcattcc cttaccaggaa 1440
gctactgtc cttgggtctg gcccggaaaca aggccggct 1500
gctaccggaa gcagccgtgg agcctggta agtcgtcat 1560
gcattgaaga ctattccac catctggagc gagagctcgc 1620
tggaggaagg cgggaaggat gcccggggtc tgctatccgg 1680
ccctgagctg gcccgtccac ggggacgggc cccagcaccc 1740
gggcccggcat tcacacctcg ggctccctca gctcccgctt 1800
agggccccgg ggcaggcatc cccagtgccc tggttctcat 1860
gccttatacat cctcatcgcc ctcaacgtcc tgctcttcta 1920
ggacagccca cacctttagt tcctggcaca gcctggccct 1980
agacggccac agagtggcc gagatccctgg cgctgcagaa 2040
tgcacaagtg gaggcagatc ctggggccct cgctggagct 2100
cgctggagaa gctgcaccaa ggcacacacag ttcagacccc 2160
ggcccgatga cagctttcc tggagacccccc ggccacgcac 2220
atggacacac agagcctcg ccggccactgc tggcaccgtg 2280
ccggccccctc ccgcacggccc aaccaggggc tgtgcagacg 2340
tgcacttttag accagggagc tggcccgccct tctggcaggc 2400
ccggctgagg ttgtgggggg ccgcctccctgg ggtgcacgat 2460
tgcattatataat ttatgggggg ccgcacagtgc cccaaataaaag 2520
ggtcagaagtaaaaaaa 2546

<210> 70
<211> 1845
<212> DNA
<213> *Homo sapien*

<220>
<221> misc_feature
<223> Incyte ID No: 5052680CB1

<400> 70

ctctggcgga ctgcctggcg gaagcgggaa cgtcgcatcc tgaggtaaag gtgcacggca 60

tcctgggaca ttagtctgg cccgggctcg gacgccccct cggatgaatg ggaccgaagc 120
 tgactcgaaa ctacagcttc ttggcagcgt cggtgttgcg cggggagaaa ggggagacccg 180
 cggcgcccc cagtgagacg ggcttccag gacgggtcgta tggctgtgcg akgcaagagg 240
 caggaggccc gcttcctggg gtacgggtac aggccggcgc ttactctgtg cgcttgcctc 300
 cccaaccctg caccggccat gcgcggggcc ttggcggtgg gcctgggttt cgcaaggctgc 360
 tgcagtaacg tgatcttc agagctccctg gcccggaaagc atccaggatg tgggaacatt 420
 gtgacatttgc cacaattttt atttattgt gtggaaaggct tcctcttgc agctgatgg 480
 ggaaggaaagc caccgctat cccaataagg tactatgcca taatgggtac catgttcttc 540
 accgtgagcg tggtaacaa ctatgcccctg aatctcaaca ttggccatgcc cctgcataatg 600
 atatttatgt cgggtctctt aattggcaac atgattttagt gaattatcat tttgaagaaa 660
 agatacagata tattcaaata tacctccatt gcccgtgtt ctgtggggat atttatttgc 720
 acttttatgt cagcaaaagca ggtgacttcc cagtcacgt tgagtggaaatgatggatc 780
 caggcatttgc tgggtgggtt acttaggtt ggggcattga ctttgtctc tctgtatgtca 840
 gcaaggatgg ggatattcca agagactctc tacaaacgtat ttgggaaaca ctccaaggag 900
 gcttgggttataatcaccgc ctttccactt ccgggttgc tttcttgc ttctgtatatt 960
 tatgaccatg cagttctatt caataagtct gagttatgt aaattcccgat catcgaggatg 1020
 accctgccc tcatgtgggtt ctacctccctc atgaacatca tcactcagta cgtgtgcattc 1080
 cgggggtgtt ttatcctcac cacggaaatgc gcctccctca ccgtcacgtc cgtcgtgacc 1140
 ctacgcaaat ttgtgagcctt catctttccatc atcttgcact tccagaacccc ctccacccctg 1200
 tggcactggc tgggcacccctt gtttgcatttgc attggggacat taatgtacac agaggtgtgg 1260
 aacaacccatgggaccacaaa aagtggacccat cagaaggacca gcaagaagaa ctgaggccctg 1320
 tctggagttac gtagaccatgt gtcgtcgatggggtggacc ctgtggatgtt ctgaccacccg 1380
 ttgcgttgcgtt gtaatggccg agcttcccgcc agtgcgtgac cagccgtgcgaaatgc 1440
 ttccaggaggg gacttccatc gttgcctcaga ctgacacatg tagactaaat agaaaacccct 1500
 cagccctaaa atagaaaaaaa gaacaggatc agactattgtt acaggcgattt catttatttc 1560
 tgggggtttt accgcacat tcaatgttattat tcttgcatttgc tgaattctgttgc 1620
 caaaaaggctt actatccatg gtcttggaaa agattgtccc ttcttgcgtt gttgtatgc 1680
 atctctgttca cagcgttcat tgcgtatccctg tgacccatttgc cttttgtctt cccatcacct 1740
 ttcccttccatc cagcgttcat tgcgtatccctg tgacccatttgc cttttgtctt cccatcacct 1800
 agtggccaaatg taaaccgaga caccagtaac agcagggtgg acttc 1845

<210> 71
 <211> 1940
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5373575CB1

<400> 71
 gtctcctgtc aaaagccatg ctcggcaggctt cgggttaccg ggcgtgtccc ctgggtgatt 60
 ttgaccgtt ccagcgtcg agcttcggctt ttctggctc gcagaaggcc tgcttgcctcc 120
 cggagcccccc cggcggtggg acagggggccg atgtacccca gagctggccc tcctgcctct 180
 gtcatggctt catcagtttc ctgggggttct tgctgtgtt ggtcacccctc cccatttctg 240
 gctgggttgc cctgaagatt gtggccacccatc acggccggat gattgtgttgc cgcctggggcc 300
 ggatccgcac ccccccaggaa cctggcatgg ttctgtcttgc ccccttccatc gactcccttc 360
 agaggggtggc tctggggaca cagcccttc acgtccctcc ctgcaagctg gcctctaagg 420
 acggggctgt gctgtccgtg ggagccgtatgg tccaggatgg catctggac cccgtgtgttgc 480
 cgggtatgac tggaaaagac ctgaaacacag ccacacgttcat gacagcccaag aacgcacatgt 540
 ccaaggccct gctcaagagg cccgtgcggg agatccatgatgg gggaaagctc aagatcagcg 600
 accagcttc tctggggatc aacgtatgttgc ccaggcccttgc ggggctggggatgg gtagaccgc 660
 tggagctggc agtggggggcc gttgtccatc cggcccccaggatc cagcccttgc gggcccaacc 720
 tggacacac cctccacgttgc ctggcccttc acttcttgc aggaaggatgtt aactcaatgg 780
 caggaggtgc cccgtcccccc gggccaggatc acaccgttgc gatgggttgc gaaatgttgc 840
 cacctgccttcc tcaagttgttgc cccaggatc gttccaaatgc gcctctggcc gggggctac 900
 tgactgtctt acaggcccttc ctgtctgttgc ccctggatc ccaaggccggg gcctgttgcacc 960
 agttcaatgtt cgtccctgtcc acggccacccaaatgccttgc cttccctggac ctcactacatg 1020
 gacggggaaatg agtggggacac ggggtggccctt atggcatccc tgatgtgggtt gttggatgtgg 1080
 cccggggcacaatgc cctggggcccttgc tggctatgttgc gggggccatgc gggccatgc 1140
 gtggacggctt gaaatgttgc gggccacccatgc tggctatgttgc gggggccatgc gggccatgc 1200
 gggcccttgc gttggatgttgc cccaggatc ccaaggccggg gggccatgc gggccatgc 1260
 gggccatgc gggccatgc cccaggatc ccaaggccggg gggccatgc gggccatgc 1320

acaatattca ggcccagcca agagcccattt aatggaggct gcaggaggct gagtccggct 1380
 gccatgcacg tctccctac aytggttctc tggacaaggc tttgtccatc ccggccccca 1440
 gctgagtgcc cagcgcttag ctgggtgcac ggtgtgattc caggaggaga gccaggcctg 1500
 ccctgcctg ctggcttccct gactggagag acaggaccca cagaaacagc ctgacagcag 1560
 ctggtttgtt ccttgtgtga gggaccaagc atgtggccca ggctctaagc tctgcgggaa 1620
 ttggagaggg atggggaggg aagggaaggc agcttcaaga agaggcttctt gtggcaagt 1680
 tacctgggta tcctggctgg cccacctcc tggctgcagt ccaggccctt gctggcggga 1740
 ttggcatgg gaaggagcag ggctgtc ttcctggcg ctgtcccaa agattctga 1800
 ctcatctgcc agctctgtcc tgcatgcctg gcgagctggg gcccaggcga gcatgaagga 1860
 gagccctgcg ttctgtgtt cttaccagag gtttgcacagc ctcagacaaa taaatgtgtt 1920
 gtttacaatg taaaaaaaaaa 1940

<210> 72
 <211> 880
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5524468CB1

<400> 72
 gcaagccga agggagcgcg gctaagagtccgcaccc tcacaacctg ggaaccggag 60
 agtagggggc gtcggctggc aagaaccgc cgtagcttccctt ccggcaaggc catccggc 120
 caccatgtc gcaactagagc ayaagagggt gatgttgcgaa actgcaacct gcacagagct 180
 gctctgtact gtcctgtgtt gtcggccca tgacctgggtt ggtgctgtg gggacactgc 240
 tctgcatgtc ggcgttgggg tttaggcaccc cggactccga gggtttcccg cccctgtgc 300
 tccacaactg cccctacaaa tgcgtatctgcg ctggcgcacct gctaagctgc actggcctag 360
 ggctgcagga cgtgccagcc gagttacactg ccgtactgc ggacactcgac ctgagccaca 420
 acgcgttcca ggcctgcgc cccggctgtt tggcccccctt cttccagctg cgcgcctgc 480
 acctagatcca caacgaacta gatgcgttgg gtcggccgtt cttcgtaac gccagcggcc 540
 tgaggctgtc cgtatctatca tctaacaacgt tgccggcgtt tggccgcac gacctcgacg 600
 ggctgggggc gctggagaag ctgtttctgt tcaataacccg cttggtgac ttggacgagc 660
 atgccttcca cggcctgcgc ggcgttccac atctctacccctt gggctgcaac gaactcgcc 720
 cttccatccctt cggaccacccg caccgttgcg gcccacccaa cctgttact ctggacctct 780
 cttccaaacccg gctggggacac atctccgtac ctgagctggc cgcgttgcgc gccttcctca 840
 agaacggcctt ctacctgcac gacaacacat aaaaaaaaaaa 880

<210> 73
 <211> 2403
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5944279CB1

<400> 73
 agccacctgc tggccaggta cccctccctt acctggggca gtgtctgcctt ggtggccact 60
 agagacagcc cagcctgggc catggaaagaa aaccggaccc ttggaaatcaga agcctggggc 120
 tcctctaggg ggtggctggc ccccccggag gccagaggag gcccacatcgct gtctctgtg 180
 ctgaacgcgc tggccactgtc tgccacccctt cggtagccgg accctgggtt gctgccttgg 240
 gggggcgtgg aggaggagga ggaggatggaa ggaaggagca gaaaggcctt cacagaagtc 300
 acccagacac agctgcagga ccctcacccctt tcccgaaac tggccctggcc catgcaggcc 360
 agacgggcac acaggcaaaatgcac agggaccagg tggctatgg ctctggaaact 420
 aagacggacc gatggggcgcg gctacttcgg aggttcaagg agaaaacaaa ggaaggcttg 480
 cgaaggctgc agccctgggc gtggacactg aagaggatcg gggggccagtt tggccggc 540
 acggagtcctt acctctccctt gtcgttccctt ctgttccctt ttaacgttgc ggcctctgtg 600
 ctcatggcctt gcatgttgcgtt tgggtggag ggcgttcccccc agggccctccc 660
 gggccggaca ttccttcgccc ctggccgttcc tataaccccccc actcccaaggg cctggtcacc 720
 ttggccaccc agtctttcaaa cttgttctcg ggtgggggtt accttggaaatg gtccctctc 780
 ttctatggctt ttcacccgc cccggccacgc ctggccgttca cttacctgtt ctggcccttt 840
 ggcgttggcc tcatctgcctt cttgttccatc ctgttgcgtt cgggttctgg gctgttgcag 900

acactgtctgg	cgggagtccgaa	ggctctgacc	agctacagcc	accgggtgtt	ctcgccctgg	960
gacttcggtc	tctgcgggaa	cgtccacgtg	cggctgcgcc	agcgcatacat	cttgcgtacgaa	1020
ttaaagggtgg	agctggagga	gacagtgggt	cggcccccagg	ctgcgggtcg	gacgcgtgggc	1080
cagcaagcca	gggtttgggt	ggtgccgggt	ctgctcaacc	tgctgggtgtt	cgcgctcctg	1140
ggggcagcct	tctatggcgt	ctactgggt	acggggtgca	ccgtggagct	gcaggagatg	1200
ccccctgtcc	aggagttgcc	actgctgaag	cttggggta	attaccttcc	gtccatcttc	1260
atcgctgggg	tcaattttgt	gctgccccc	gtgttcaagc	tcattgtccc	actggagggc	1320
tacactccgaa	gtcgccagat	cgtttttatac	ctgctcagga	ccgtgtttct	tcgcctcgcc	1380
tccctgggtt	tctctgtctt	ctctctctgg	aatcagatca	cttgggggg	cgactecgag	1440
gctgaggact	gcaaaaacctg	tggctacaat	tacaaaacaac	ttccgtgtcg	ggagactgtc	1500
ctggcccaagg	aaatgtacaa	acttctgtc	tttgatctgc	tgactgtctt	ggcagtcgcg	1560
ctgctcatcc	atgttcttag	aaagctcctc	tgtggcctct	gtcctggggc	gctgggtcg	1620
ctggcggggaa	cccaggagtt	ccaggtgccc	gacgagggtgc	tggggctcat	ctacgcgcag	1680
acgggtgtct	gggtggggag	tttttctgc	ccttactgc	ccctgtttaa	cacgtcaag	1740
ttcctgtgc	ttttctacct	gaagaagctt	accctcttct	ccacctgtc	cccgctgccc	1800
cgcacctcc	gggcctccgc	ggcgaattt	ttttccctt	tggctcttct	cctgggtctg	1860
gccatctcca	gcgttccccct	gctttacagc	atcttcttga	tcccgcctt	taagctttgt	1920
ggtccattcc	ggggcagtc	gtccatctgg	gccccagatcc	ctgagtttat	ttccagcctc	1980
cctgagacca	cccagaattt	cctcttcttc	ctggggaccc	aggcttttgc	tgtccccctt	2040
ctgctgtatct	ccagcatctt	gatggcgatc	actgtggctc	tggcttaactc	ctacggacgc	2100
ctcatctctg	agctcaaacg	tccagagacag	acggaggcgcc	agaataaaagt	tttcctggca	2160
cggcgcgtgt	tggcgctgac	ctccccaaaa	ccggctttt	gaccggccga	ggccacgtcc	2220
cgctttcaga	cccccaggcccc	attgttaagcc	taggtcacaa	catctgtaaa	ctagggaaac	2280
tggagaagac	tccacgcctt	tccagcttgc	gtatctggag	atttccaggg	cccctcgccg	2340
ccacgtccct	gactctcgaa	tgatcttccct	tgtatcaata	aatacagccg	aggttgggtga	2400
acc						2403

<210> 74
<211> 2850
<212> DNA
<213> *Homo sapiens*

```
<220>
<221> misc_feature
<223> Incyte ID No: 6114480CB1
```

<400> 74

```

cggctcgagc ggctcgagtg aagagccctc ccacggctcc tgccgctgag acagctggcc 60
tgacctcaa atcatccatc caccctgtc gtcatctgtt ttcatagtgt gagatcaacc 120
cacaggaata tccatggctt ttgtgctcat tttgggtctc agtttctacg agctgggtc 180
aggacagtgg caagtcactg gaccggcaa gtttgcctag gccttgggtgg gggaggacgc 240
cgtttctcc tgctccctct tcctgtgagac cagtgcagag gctatggaaag tgccgttctt 300
caggaatcag ttccatgctg tggtccacct ctacagagat gggaaagact gggaaatctaa 360
gcagatgcca cagtatcgag ggagaactga gtttgcgtgg aactccattg caggggggcgc 420
tgtctctcta aggctaaaaa acatcactcc ctggacatc gcctgtgtt ggtgtgggtt 480
cagttccccag attacgtat aggaggccac ctgggagctg cgggtggcag cactgggtc 540
acttcccttc atttccatcg tgggatatgt tgacggaggt atccagttac tctgcctgtc 600
ctcaggctgg ttccccccagc ccacagccaa gtggaaagggt ccacaaggac aggatttgc 660
ttcagactcc agagcaaattg cagatgggtt cagcctgttat gatgtggaga tctccattat 720
agtccaggaa aatgtgggaa gcatattgtt ttccatccac cttgctgagc agagtcatga 780
ggtggaatcc aaggatattga taggagagac gttttccag ccctcacctt ggcgcctggc 840
ttctattta ctcgggttac tctgtgggtc cctgtgtgtt gttgtcatgg ggatgataat 900
tgttttcttc aaatccaaag ggaaaatcca ggccgaactg gactggagaa gaaagcacgg 960
acaggcagaa ttgagagacg cccggaaaca cgcagtggag gtgactctgg atccagagac 1020
ggctcacccg aagctctgca tttctgtatc gaaaactgtt accccatagaa aagctcccc 1080
ggaggtgcct cactctgaga agagatttac aaggaagagt gtgggtggctt ctcagggttt 1140
ccaagcaggg agacattact gggaggtggc cgtgggacaa aatgttaggt ggtatgtggg 1200
agtgtgtcgg gatgacgtat acagggggaa gaacaatgtt acttgggtc ccaacaatgg 1260
gttattgggtc ctaagactg caacagaaaca tttgtattttt acattcaatc cccatattt 1320
cagcctcccc cccaggcaccc ctccatcacg agtaggggtc ttccctggact atgaggggtgg 1380
gaccatctcc ttcttcaata caaatgacca gtcccttatt tataccctgc tgacatgtca 1440
gttgaaggc ttgttgagac cctatattca gcatgcgtat tatgacgagg aaaaggggac 1500
tcccatattc atatgtccag tgcctgggg atgagacaga gaagaccctg cttaaaggggc 1560

```

cccacaccac agaccagac acagccaagg gagagtgc tc ccgacaggtg gccccagctt 1620
cctctccgga gcctgcgcac agagagtac gcccccaact ctccctttagg gagctgagg 1680
tcttctgcctt tgagccctgc agcagcggca gtcacagctt ccagatgagg ggggattggc 1740
ctgaccctgt gggagtcaaga agccatggct gcccctgaagt ggggacggaa tagactcaca 1800
tttaggtttag ttgtgaaaa ctccatccag ctaagcgatc ttgaacaagt cacaacctcc 1860
caggctccctc atttgcttagt cacggacagt gattctgc tc acaggtga agattaaaga 1920
gacaacgaat gtgaatcatg cttgcaggt tgagggcaca gtgtttgcta atgatgttt 1980
tttatattat acattttccc accataaaact ctgtttgctt attccacatt aatttacttt 2040
tctctatacc aaatcaccctt tggaaatagtt attgaacacc tgctttgtga ggctcaaaga 2100
ataaaagagga ggttaggattt ttcactgatt ctataagccc agcattacct gataccaaaa 2160
ccaggccaaag aaaacagaag aagaggaagg aaaactacag gtcctatcc ctcattaaaca 2220
cagacacaaaa aattctaaat aaaattttaa caaattaaac taaacaatat atttaaagat 2280
gatataaac tactcagtgt gtttgtccc acaaatgcag agttggtttta atatttaaat 2340
atcaaccagt gtaattcagc acattaatag agtaggaaag aaaccatagg aaggagagac 2400
gaaggaaaga gaaaaggggg cggcccaact agtgagctgg caccccccggg atttccggc 2460
gcggactctc aggggggtag cgggtatcga agcagctt aagacggaga ctcggggggg 2520
gccgggacca aattcccaa gagggggcggt tattcggcgcg cgctggcctc ggtttaaac 2580
gccggagtgg gaaccagggtt gtaccaagtt agggcttgag ccgaccgtat gttgggcttc 2640
cagctagaga ctgcaaggaa agaagaagga ggctagagggc caggatagtg gagagagaac 2700
agacgcgaga acaaggggaa atgaggacgc gaggggagag atgaggagag atagagatgt 2760
gacgagcgc gaggagagcc gggaccggag aggcaacggga gaagtgacga acgaagaaag 2820
acaaggggcgc gaaggaaacgc agaagtggaa 2850

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
22 February 2001 (22.02.2001)

PCT

(10) International Publication Number
WO 01/12662 A3

(51) International Patent Classification⁷: **C12N 15/12.**
C07K 14/47, 14/705, 16/18, 16/28, C12Q 1/68, A61K
38/17, A01K 67/027, G01N 33/50

(21) International Application Number: PCT/US00/22315

(22) International Filing Date: 14 August 2000 (14.08.2000)

(25) Filing Language: English

Francisco, CA 94122 (US). **AZIMZAI, Yalda** [US/US];
2045 Rock Springs Drive, Hayward, CA 94545 (US).
BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road,
San Leandro, CA 94577 (US). **LU, Dyung, Aina, M.**
[US/US]; 55 Park Belmont Place, San Jose, CA 95136
(US). **PATTERSON, Chandra** [US/US]; 490 Sherwood
Way #1, Menlo Park, CA 94025 (US).

(26) Publication Language: English

(74) Agents: **HAMLET-COX, Diana** et al.: Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(30) Priority Data:
60/149,641 17 August 1999 (17.08.1999) US
60/164,203 9 November 1999 (09.11.1999) US

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier applications:

US 60/149,641 (CIP)
Filed on 17 August 1999 (17.08.1999)
US 60/164,203 (CIP)
Filed on 9 November 1999 (09.11.1999)

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (for all designated States except US): **INCYTE
GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

Published:

— with international search report

(72) Inventors; and

(88) Date of publication of the international search report:
6 December 2001

(75) Inventors/Applicants (for US only): **LAL, Preeti**
[IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US).
YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA
94087 (US). **TANG, Y., Tom** [CN/US]; 4230 Ranwick
Court, San Jose, CA 95118 (US). **BANDMAN, Olga**
[US/US]; 366 Anna Avenue, Mountain View, CA 94043
(US). **BURFORD, Neil** [GB/US]; 1308 4th Avenue, San

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/12662 A3

(54) Title: MEMBRANE ASSOCIATED PROTEINS

(57) Abstract: The invention provides human membrane associated proteins (MEMAP) and polynucleotides which identify and encode MEMAP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of MEMAP.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/22315

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7	C12N15/12	C07K14/47	C07K14/705	C07K16/18	C07K16/28
	C12Q1/68	A61K38/17	A01K67/027	G01N33/50	

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K C12Q A61K A01K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, EMBL, STRAND, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL [Online] EMBL; ID AA984180, AC A984180, 28 May 1998 (1998-05-28) HILLIER L ET AL.: "am82a04.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone IMAGE:1629582 3', mRNA sequence" XP002159906 Note: 99.6 % nt seq identity with SEQ ID NO:1 in 500 nt overlap. the whole document</p> <p>---</p> <p style="text-align: center;">-/-</p>	1-19,22, 25-28

 Further documents are listed in the continuation of box C. Patent family members are listed in annex.

° Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

13 February 2001

Date of mailing of the international search report

15.05.01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

van de Kamp, M

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/22315

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL [Online] EMBL; ID AI243841, AC AI243841, 5 November 1998 (1998-11-05) STRAUSBERG R: "qh89e04.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone IMAGE:1854I74 3' similar to contains element MER26 repetitive element ;, mRNA sequence" XP002159907 Note: 100.0 % nt seq identity with SEQ ID N0:1 in 412 nt overlap. the whole document</p> <p>---</p>	1-19,22, 25-28
X	<p>DATABASE EMBL [Online] embl; ID AC005924, AC AC005924, 5 November 1998 (1998-11-05) MUZNY D ET AL.: "Homo sapiens Chr.14 PAC RPCI4-794B2 (Roswell Park Cancer Institute Human PAC Library) complete sequence" XP002159908 Note: 93.4 % nt seq identity of nt 17917-19843 with SEQ ID N0:1 in 1932 nt overlap. page 13</p> <p>---</p>	3-5, 11-15
A	<p>DALMAU J ET AL.: "Mal, a novel neuron- and testis-specific protein, is recognized by the serum of patients with paraneoplastic neurological disorders" BRAIN, vol. 122, no. 1, January 1999 (1999-01), pages 27-39, XP002159905 Note: 54.7 % aa seq identity of Mal with SEQ ID N0:38 in 333 aa overlap. abstract page 33, left-hand column, line 17 -page 34, left-hand column, line 2; figure 5 page 37, left-hand column, line 9-15</p> <p>---</p>	1-7,9-15
A	<p>WO 98 21328 A (KATO SEISHI ;PROTEGENE INC (JP); SEKINE SHINGO (JP); SAGAMI CHEM R) 22 May 1998 (1998-05-22) the whole document</p> <p>---</p>	
A	<p>WO 98 22491 A (MILLENNIUM BIOTHERAPEUTICS INC) 28 May 1998 (1998-05-28) the whole document</p> <p>---</p> <p>-/-</p>	

INTERNATIONAL SEARCH REPORT

Inter	inal Application No
PCT/US 00/22315	

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>YOKOYAMA-KOBAYASHI M ET AL.: "Selection of cDNAs encoding putative type II membrane proteins on the cell surface from a human full-length cDNA bank" GENE, vol. 228, no. 1-2, 4 March 1999 (1999-03-04), pages 161-167, XP004159146 the whole document</p> <p>-----</p>	

INTERNATIONAL SEARCH REPORT

Ir. International application No.
PCT/US 00/22315

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

Remark (1): Although claims 18, 21 and 24 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition. Concerning claims 21 and 24, see also Remark (2).

2. Claims Nos.: 20, 21, 23, 24 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-28 all partially

Remark on Protest

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-28 (all partially)

An isolated polypeptide comprising or consisting of the amino acid sequence of SEQ ID NO:1 or homologs or fragments thereof, an isolated polynucleotide encoding a polypeptide as said, a recombinant polynucleotide containing a polynucleotide as said linked to a promoter, a host cell and a transgenic organism containing a recombinant polynucleotide as said, a method for producing a polypeptide as said, an antibody binding to a polypeptide as said, an isolated polynucleotide comprising a nucleotide sequence of SEQ ID NO:38 or homologs or fragments thereof, methods of detecting a target polynucleotide having a nucleotide sequence of a polynucleotide as said, a pharmaceutical composition comprising a polypeptide as said and the use of a composition as said in medical treatment, methods for screening for a compound that is effective as an agonist or an antagonist of a polypeptide as said, and compositions comprising a agonist or an antagonist as said as well as the use of compositions as said in medical treatment, methods for screening for a compound that binds to a polypeptide as said, that modulates the activity of a polypeptide as said, or that alters the expression of a polynucleotide as said, and a method for assessing the toxicity of a test compound involving the detection of a polynucleotide as said.

2. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOs 2 and 39.

3. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOs 3 and 40.

4. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOs 4 and 41.

5. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOs 5 and 42.

6. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOs 6 and 43.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

7. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 7 and 44.

8. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 8 and 45.

9. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 9 and 46.

10. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 10 and 47.

11. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 11 and 48.

12. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 12 and 49.

13. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 13 and 50.

14. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 14 and 51.

15. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 15 and 52.

16. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 16 and 53.

17. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 17 and 54.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

18. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 18 and 55.

19. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 19 and 56.

20. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 20 and 57.

21. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 21 and 58.

22. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 22 and 59.

23. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 23 and 60.

24. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 24 and 61.

25. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 25 and 62.

26. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 26 and 63.

27. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 27 and 64.

28. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID NOS 28 and 65.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

29. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 29 and 66.

30. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 30 and 67.

31. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 31 and 68.

32. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 32 and 69.

33. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 33 and 70.

34. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 34 and 71.

35. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 35 and 72.

36. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 36 and 73.

37. Claims: 1-28 (all partially)

As invention 1, but concerning SEQ ID N0s 37 and 74.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 20, 21, 23, 24

Remark (2): Claims 20, 21, 23 and 24 refer to agonists and antagonists of the polypeptide, and to compositions comprising such agonists and antagonists, without giving a true technical characterisation. Moreover, no specific compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT). No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the results to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 00/22315

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9821328 A	22-05-1998	AU 4885297 A		03-06-1998
		EP 0941320 A		15-09-1999
WO 9822491 A	28-05-1998	US 5952171 A		14-09-1999
		AU 5354798 A		10-06-1998
		EP 0948519 A		13-10-1999